CORRECTED VERSION*

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, 15/85, 5/10, C07K 14/705, 16/28, C12Q 1/68, A61K 38/17, 48/00

A3

(11) International Publication Number: **WO 98/31799**

(43) International Publication Date: 23 July 1998 (23.07.98)

(21) International Application Number: PCT/US98/00959

(22) International Filing Date: 21 January 1998 (21.01.98)

(30) Priority Data:

60/034,204	21 January 1997 (21.01.97)	US
60/034,205	21 January 1997 (21.01.97)	US

(71) Applicant *(for all designated States except US)*: HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: NI, Jian [CN/US]; 5502 Manorfield Road, Rockville, MD 20853 (US). GENTZ, Reiner, L. [DE/US]; 13404 Fairland Park Drive, Silver Spring, MD 20904 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US).

(74) Agents: BROOKES, A., Anders et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
5 November 1998 (05.11.98)

(54) Title: POLYNUCLEOTIDES AND POLYPEPTIDES ENCODING RECEPTORS

(57) Abstract

Receptor polypeptides and polynucleotides and methods for producing such polypeptides by recombinant techniques are disclosed. Also disclosed are methods for utilizing receptor polypeptides and polynucleotides in the design of protocols for the treatment of diseases and diagnostic assays for such conditions.

*(Referred to in PCT Gazette No. 49/1998, Section II)

BNSDOCID: <WO___9831799A3_IA>

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

POLYNUCLEOTIDES AND POLYPEPTIDES ENCODING RECEPTORS

FIELD OF INVENTION

This invention relates to newly identified polynucleotides and the polypeptides encoded by them, the use of such polynucleotides and polypeptides, and their production. More particularly, the polynucleotides and polypeptides of the present invention relate to specific receptor families described in the specification and known in the art. The invention also relates to inhibiting or activating the action of such polynucleotides and polypeptides.

BACKGROUND OF THE INVENTION

Receptor proteins are found on the membrane of the cells and are generally involved in signal transduction. There are many types of receptor proteins, and for convenience, these proteins are grouped in families based on similarity in structure and function.

For example, the TM4SF superfamily of cell surface proteins, also known as the tetraspan receptor superfamily, is comprised of at least seventeen individual gene products (these include CD9, CD20, CD37, CD53, CD63, CD81, CD82, A15, CO-029, Sm23, RDS, Uro B, Uro A, SAS, Rom-1, PETA3, and YKK8). The TM4SF superfamily is the second largest group in the CD antigen superfamily. Each member of the TM4SF superfamily can be characterized by several putative physical features including four highly conserved transmembrane domains, two divergent extracellular loops, and two short and highly divergent cytoplasmic tails. Expression patterns for members of the TM4SF superfamily tend to be rather broad and can vary widely between members. The functional roles of TM4SF superfamily members are primarily associated with signal transduction events and pathways, but also include cell adhesion in platelets and other lymphocytic and non-lymphocytic cell lines, as well as cell motility, proliferation, and metastasis. In addition, recent evidence suggests that a subset of the members of the TM4SF superfamily may function as potassium channel molecules.

One member of the TM4SF family, CD20, is a four membrane spanning domain cell surface phosphoprotein expressed exclusively on B lymphocytes. Although the precise functional role of CD20 has yet to be determined, it is thought to function primarily as a receptor during B-cell activation. Furthermore, a large number of experimental observations suggest several additional speculative roles for the CD20 molecule. For example, CD20-specific immunoprecipitation of biochemically cross-linked plasma membrane proteins suggests that CD20 assumes a multimeric structural

conformation characteristic of other previously described membrane channel proteins. Further experimentation has revealed that expression of exogenous CD20 on the cell surface specifically increases Ca^{2+} conductance across the plasma membrane. Together, these results suggest that CD20 complexes may function as B-cell specific Ca^{2+} ion channels. In addition, monoclonal antibodies raised against CD20 have been used to stimulate resting B-cells to transition out of the G0/G1 segment of the cell cycle. It has also been demonstrated that CD20 is associated with both serine and tyrosine kinases and, more specifically, that CD20 is associated, although not directly, with the Src family of tyrosine kinases including p56/53lyn, p56lck, and p59fyn.

A second example of a receptor subfamily, called sialoadhesin molecules, belongs to the Ig superfamily of receptor-like molecules. The more than 100 members of the Ig superfamily are generally considered to engage in specific cell-cell interactions through which intercellular communication may occur. In addition to classical protein-protein interactions, intercellular communication may also be mediated through protein-carbohydrate interactions. In fact, all members of the sialoadhesin family of the Ig superfamily are capable of mediating protein-sialic acid binding interactions. To date, only a small number of proteins have been assigned to the sialoadhesin family including sialoadhesin, CD33, CD22, the myelin-associated glycoprotein (MAG), and the Schwann cell myelin protein (SMP). Each of these proteins is expressed in a restricted subset of cell types. For example, CD22 and CD33 are expressed exclusively by B-lymphocytes and cells of the myelomonocytic lineage, respectively.

Similarly, galectins are a family of the lectin superfamily of carbohydrate-binding proteins which have a high affinity for b-galactoside sugars. Although a large number of glycoproteins containing b-galactoside sugars are produced by the cell, only a few will bind to known galectins *in vitro*. Such apparent binding specificity suggests a highly specific functional role for the galectins. Galectin 1 (conventionally termed *LGALS1* for lectin, galactoside-binding, soluble -1) is thought to specifically bind laminin, a highly polylactosaminated cellular glycoprotein, as well as the highly polylactosaminated lysosome-associated membrane proteins (LAMPs). Galectin 1 has also been shown to bind specifically to a lactosamine-containing glycolipid found on olfactory neurons and to integrin a_7b_1 on skeletal muscle cells. Galectin 3 has also been observed to bind specifically to laminin, immunoglobulin E and its receptor, and bacterial lipopolysaccharides.

Various galectins have been shown to function in the mechanisms of intercellular communication. For example, depending on cell type, galectin 1 has been observed to modulate cell adhesion either positively or negatively. More specifically, galectin 1 appears to inhibit cell adhesion of skeletal muscle presumably by galectin 1-mediated disruption of laminin-integrin a_7b_1 interactions. Alternatively, galectin 1 appears to promote cell adhesion in several non-skeletal muscle cell types examined

presumably by a glycoconjugate cross-linking mechanism. Galectin 3 has also been observed to function in modulating cell-adhesion, as well as in the activation of certain immune cells by cross-linking IgE and IgE receptors. In addition, galectins have been observed to be involved in the regulation of immune cell activity, as well as in such diverse processes as cell adhesion, proliferation, inflammation, autoimmunity, and metastasis of tumor cells. Furthermore, a galectin-like antigen designated HOM-HD-21 was recently found to be highly expressed in a Hodgkin's Disease cDNA library. Very recently, a novel galectin, termed PCTA-1, was identified as a specific cell surface marker on human prostate cancer cell lines and patient-derived carcinomas. Galectins have also been found to function intracellularly as a component of ribonucleoprotein complexes. Finally, galectins 1 and 3 have each been found to modulate T-cell growth and apoptosis by interaction with CD45 and possibly Bcl2, respectively.

A relatively new family of cell-surface proteins has been identified and termed the Ly6 superfamily. The members of this family include murine and human SCA-2, rat Ly-6 (also termed ThB), human CD59 [also known as protectin or membrane attack complex inhibition factor (MACIF)], and E48 antigen. The determination of an initial functional role for SCA-2 may lie in an analysis of its expression profile with regard to the complex process of hematopoiesis. SCA-2 is highly expressed in early thymic precusor cells. In turn, progeny of the intrathymic precusor population continue to express SCA-2, but only until the point of transition occurs from blast cell to small cell. Further experimental evidence demonstrates that mature thymocytes and peripheral T-cells do not express detectable levels of SCA-2, whereas mature, peripheral B-cells do continue to express SCA-2. As a result, it seems very likely that SCA-2 plays an important role in thymocyte maturation and differentiation. A plausible explanation for this functional hypothesis is that SCA-2 may act as a receptor for a unknown cytokine which regulates thymocyte maturation and differentiation.

In addition, CD59 is a recently identified integral membrane protein which appears to be involved in the regulation of complement. Recent studies show that the CD59 antigen may prevent damage from complement C5b-9 and protect astrocytes during inflammatory and infectious disorders of the nervous system. Expression of recombinant human CD59 on porcine donor organs have been shown to prevent complement-mediated lysis and activation of endothelial cells that leads to hyperacute rejection. Recently, researchers at Alexion Pharmaceuticals (New Haven, CT) reported on the production of transgenic pigs which expressed human CD59. In these animals, xenogeneic organs were resistant to hyperacute rejection. (Fodor, et al., "Expression of a functional human complement inhibitor in a transgenic pig as a model for the prevention of xenogeneic hyperacute organ rejection," Proc. Natl. Acad. Sci., 91:1153-11157 (1994).) The same company also reported that expression of recombinant transmembrane CD59 in paroxysmal nocturnal hemoglobinuria (PNH) B-

cells confers resistance to human complement. (Rother et al., "Expression of recombinant transmembrane CD59 in paroxysmal nocturnal hemoglobinuria B-cells confers resistance to human complement," Blood, 84:2604-2611 (1994).) PNH is an acquired hematopoietic disorder characterized by complement-mediated hemolytic anemia, pancytopenia, and venous thrombosis. It is thought that retroviral gene therapy with this molecule could provide a treatment for PNH patients.

A final Ly6 superfamily member, the E48 antigen, is involved in intercellular adhesion between keratinocyte cells of the squamous epithelium. Such keratinocytes are attached to adjoining cells by large numbers of desmosomes, which are thought to play a role in the transition of transformed keratinocytes to metastatic tumor cells. Treatment with a monoclonal antibody raised against the E48 antigen has been successful in the eradication of residual, postoperative squamous cell carcinoma cells of the upper aerodigestive tract in several *in vivo* models and, to some degree, in humans. (van Dongen, et al., "Progress in radioimmunotherapy of head and neck cancer," Oncol. Rep. 1:259-264 (1994).) The gene encoding the E48 antigen has been mapped to the q24-qter region of human chromosome 8. Interestingly, a number of human diseases have been mapped to this region of chromosome 8 including Langer-Giedion syndrome, brachio-otorhinolaryngeal syndrome, trichorhinolaryngeal syndrome, and epidermolysis bullosa simplex.

A further example of a receptor family includes the prohibitin receptors. The prohibitin gene product is expressed in a wide variety of tissues and has been implicated as a component of a number of anti-proliferative mechanisms. The prohibitin gene encodes a 30 kD postsynthetically modified polypeptide located primarily in the mitochondria, but also may be associated with the IgM receptor on the B-cell plasma membrane. The protein functionally inhibits DNA synthesis and entry into S phase of the cell cycle by an unknown mechanism. Interestingly, although the prohibitin gene product is hypothesized to be involved in the maintenance of senescence and the prevention of cancer, one study found that, although somatic mutations in the prohibitin gene were present in a small number of breast cancers, no mutations were identified in any other breast, ovary, liver, and lung cancers examined. (Sato *et al.*, Genomics 17:762-764 (1993).) However, the prohibitin gene has been mapped to human chromosome 17q12-21, the same region thought to contain the gene involved in sporadic breast cancer. Furthermore, DNA sequence analysis of the prohibitin gene identified somatic mutation in 4 of 23 cases of sporadic breast cancer examined. Thus, prohibitin family members may be involved in the development of cancer.

Moreover, the EGFR family of plasma membrane proteins are an integral component of normal cellular proliferation and in the pathogenesis of the cancerous state. The family is relatively small and includes the EGFR, c-erbB-2, c-erbB-3, and others. Various cancers are correlated with aberrant expression of one or more of these

genes. A number of ligands have been identified which bind to the EGFR-like receptors listed above including TGF-a, heparin-binding EGF, amphiregulin, criptoregulin, heregulin, and others. A large fraction of adenocarcinomas examined to date, especially those of the breast, colon, and pancreas, are typified by the amplification or overexpression of the c-erbB-2 gene. EGF, or an analogous ligand, initiates the cellular growth factor response by binding to the EGFR, or EGFR-related, receptor. Following the binding event, the receptor molecule dimerizes activating its intracellular tyrosine kinase domain. This event results in the phosphorylation of specific tyrosine residues near the carboxy terminus of the receptor. The diversity of signals able to be transduced through the relatively small number of EGFR-related receptor molecules is amplified considerably by the recent finding that EGFR-like receptor molecules can function when dimerized with other EGFR family members forming heterodimers.

Members of the EGFR-related family of integral membrane proteins have been implicated in the pathogenesis of a number of human disease-states. For example, a mutation in the EGFR itself appears to play an important role in the development of glioblastomas. (Sang et al., J. Neurosurg 82:841-846 (1995).) The EGFR gene is amplified or overexpressed in the majority of primary human glioblastomas. Although not conferring a distinct advantage on cell growth, an increase in EGFR expression was found to confer an increase in the ability of glioma cells to maintain anchorage-independent growth in soft agar especially in response to EGF and retinoic acid. Anchorage-independent growth *in vitro* correlates highly with tumorigenicity *in vivo*, therefore, it is likely that cells which express abnormally high levels of EGFR in human glioblastoma cells may be involved in the high potential for these cells to cause tumors *in vivo*.

Moreover, overexpression or amplification of c-erbB-2 has been reported to be involved in a high number adenocarcinomas, particularly of the breast, colon, and pancreas, and in a small proportion of ovarian carcinomas.

Thus, there is a clear need for identifying and exploiting novel members of the receptor families, such as those described above. Although structurally related, these receptors will likely possess diverse and multifaceted functions in a variety of cell and tissue types. Receptor type molecules should prove useful in target based screens for small molecules and other such pharmacologically valuable factors. Monoclonal antibodies raised against such receptors may prove useful as therapeutics in an anti-tumor, diagnostic, or other capacity. Furthermore, receptors described here may prove useful in an active or passive immunotherapeutical role in patients with cancer or other immunocompromised disease states.

SUMMARY OF THE INVENTION

In one aspect, the invention relates to receptor polypeptides and polynucleotides, as well as the methods for their production. Another aspect of the invention relates to methods for using such receptor polypeptides and polynucleotides. Such uses include the treatment of the specified diseases, among others. In still another aspect, the invention relates to methods to identify agonists and antagonists using the materials provided by the invention, and treating conditions associated with receptor imbalance with the identified compounds. Yet another aspect of the invention relates to diagnostic assays for detecting diseases associated with inappropriate receptor activity or levels.

DESCRIPTION OF THE INVENTION

Definitions

The following definitions are provided to facilitate understanding of certain terms used frequently herein.

"Receptor" refers, among others, to a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:Y, or an allelic variant thereof.

"Receptor Activity" or "Biological Activity of the Receptor" refers to the metabolic or physiologic function of said receptor including similar activities or improved activities or these activities with decreased undesirable side-effects. Also included are antigenic and immunogenic activities of said receptor.

"Receptor gene" refers to a polynucleotide comprising the nucleotide sequence set forth in SEQ ID NO:X or allelic variants thereof and/or their complements.

"SEQ ID NO:X" comprises all or a substantial portion of the polynucleotide encoding each receptor of the invention. The value X for the nucleotide sequence is an integer specified in Table 1. This nucleotide sequence was translated into the receptor polypeptide identified in Table 1 as "SEQ ID NO:Y," where the value of Y for each receptor polypeptide is an integer defined in Table 1.

The invention further provides a composition of matter comprising a nucleic acid molecule which comprises a human cDNA clone identified by a cDNA Clone ID (Identifier) in Table 1, which DNA molecule is contained in the material deposited with the American Type Culture Collection ("ATCC") and given the ATCC Deposit Number shown in Table 1 for that cDNA clone. The ATCC is located at American Type Culture Collection (ATCC), 12301 Park Lawn Drive, Rockville, Maryland 20852, USA. The deposit has been made under the terms of the Budapest Treaty on the international recognition of the deposit of micro-organisms for purposes of patent procedure. The strain will be irrevocably and without restriction or condition released to the public upon the issuance of a patent. The deposit is provided merely as convenience to those of skill in the art and is not an admission that a deposit is required for enablement, such as that

required under 35 U.S.C. §112. The nucleotide sequence of the polynucleotides contained in the deposited material, as well as the amino acid sequence of the polypeptide encoded thereby, are controlling in the event of any conflict with any description of sequences herein.

"Antibodies" as used herein includes polyclonal and monoclonal antibodies, chimeric, single chain, and humanized antibodies, as well as Fab fragments, including the products of an Fab or other immunoglobulin expression library.

"Isolated" means altered "by the hand of man" from the natural state. If an "isolated" composition or substance occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living animal is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein.

"Polynucleotide" generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. "Polynucleotides" include, without limitation single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, "polynucleotide" refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The term polynucleotide also includes DNAs or RNAs containing one or more modified bases and DNAs or RNAs with backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications has been made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically or metabolically modified forms of polynucleotides as typically found in nature, as well as the chemical forms of DNA and RNA characteristic of viruses and cells. "Polynucleotide" also embraces relatively short polynucleotides, often referred to as oligonucleotides.

"Polypeptide" refers to any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres. "Polypeptide" refers to both short chains, commonly referred to as peptides, oligopeptides or oligomers, and to longer chains, generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene-encoded amino acids. "Polypeptides" include amino acid sequences modified either by natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can

occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York, 1993 and Wold, F., Posttranslational Protein Modifications: Perspectives and Prospects, pgs. 1-12 in POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, 1983; Seifter et al., "Analysis for protein modifications and nonprotein cofactors", Meth Enzymol (1990) 182:626-646 and Rattan et al., "Protein Synthesis: Posttranslational Modifications and Aging", Ann NY Acad Sci (1992) 663:48-62.)

"Variant" as the term is used herein, is a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide respectively, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to

occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis.

"Identity" is a measure of the identity of nucleotide sequences or amino acid sequences. In general, the sequences are aligned so that the highest order match is obtained. "Identity" per se has an art-recognized meaning and can be calculated using published techniques. (See, e.g.: COMPUTATIONAL MOLECULAR BIOLOGY, Lesk, A.M., ed., Oxford University Press, New York, 1988; BIOCOMPUTING: INFORMATICS AND GENOME PROJECTS, Smith, D.W., ed., Academic Press, New York, 1993; COMPUTER ANALYSIS OF SEQUENCE DATA, PART I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; SEQUENCE ANALYSIS IN MOLECULAR BIOLOGY, von Heinje, G., Academic Press, 1987; and SEQUENCE ANALYSIS PRIMER, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991.) While there exist a number of methods to measure identity between two polynucleotide or polypeptide sequences, the term "identity" is well known to skilled artisans. (Carillo, H., and Lipton, D., SIAM J Applied Math (1988) 48:1073.) Methods commonly employed to determine identity or similarity between two sequences include, but are not limited to, those disclosed in Guide to Huge Computers, Martin J. Bishop, ed., Academic Press, San Diego, 1994, and Carillo, H., and Lipton, D., SIAM J Applied Math (1988) 48:1073. Methods to determine identity and similarity are codified in computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, GCS program package (Devereux, J., et al., Nucleic Acids Research (1984) 12(1):387), BLASTP, BLASTN, FASTA (Atschul, S.F. et al., J Molec Biol (1990) 215:403.)

As an illustration, by a polynucleotide having a nucleotide sequence having at least, for example, 95% "identity" to a reference nucleotide sequence of SEQ ID NO:X is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence of SEQ ID NO: X. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. These mutations of the reference sequence may occur at the 5 ' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence.

Similarly, by a polypeptide having an amino acid sequence having at least, for example, 95% "identity" to a reference amino acid sequence of SEQ ID NO:Y is intended that the amino acid sequence of the polypeptide is identical to the reference sequence except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the reference amino acid of SEQ ID NO:Y. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

Polypeptides of the Invention

In one aspect, the present invention relates to receptor polypeptides (or receptor proteins). The receptor polypeptides include the polypeptide of SEQ ID NO:Y; as well as polypeptides comprising the amino acid sequence of SEQ ID NO:Y; and polypeptides comprising the amino acid sequence which have at least 80% identity to that of SEQ ID NO:Y over its entire length, and still more preferably at least 90% identity, and even still more preferably at least 95% identity to SEQ ID NO:Y. Furthermore, those with at least 97-99% identity to SEQ ID NO:Y are highly preferred. Also included within receptor polypeptides are polypeptides having the amino acid sequence which have at least 80% identity to the polypeptide having the amino acid sequence of SEQ ID NO:Y over its entire length, and still more preferably at least 90% identity, and even still more preferably at least 95% identity to SEQ ID NO:Y. Furthermore, those with at least 97-99% are highly preferred. Preferably receptor polypeptides exhibit at least one biological activity of the receptor.

The receptor polypeptides may be in the form of the "mature" protein or may be a part of a larger protein such as a fusion protein. It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification such as multiple histidine residues, or an additional sequence for stability during recombinant production.

Fragments of the receptor polypeptides are also included in the invention. A "fragment" is a polypeptide having an amino acid sequence that entirely is the same as part, but not all, of the amino acid sequence of the aforementioned receptor polypeptides. As with receptor polypeptides, fragments may be "free-standing," or comprised within a larger polypeptide of which they form a part or region, most

preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, and 101 to the end of receptor polypeptide. In this context "about" includes the particularly recited ranges larger or smaller by several, 5, 4, 3, 2 or 1 amino acid at either extreme or at both extremes.

Preferred fragments include, for example, truncation polypeptides having the amino acid sequence of receptor polypeptides, except for deletion of a continuous series of residues that includes the amino terminus, or a continuous series of residues that includes the carboxyl terminus or deletion of two continuous series of residues, one including the amino terminus and one including the carboxyl terminus.

Also preferred are fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. The "domains" of each receptor polypeptide are illustrated in the Figures. The Figures compare SEQ ID NO:Y to the closest know homologue. Identical amino acids shared between the two polypeptides are shaded, while conservative amino acid changes are boxed. By examining the regions or amino acids shaded and/or boxed, the skilled artisan can readily identify conserved domains between the two polypeptides. The amino acids sequences of SEQ ID NO:Y falling within these conserved domains are "fragments" and are specifically contemplated by the present invention. Especially preferred is the extracellular domains of a receptor of the invention. Soluble extracellular domains have antagonist activity mediated by competition with a receptor ligand.

Other preferred fragments are biologically active fragments. Biologically active fragments are those that mediate receptor activity, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Also included are those that are antigenic or immunogenic in an animal, especially in a human.

Preferably, all of these polypeptide fragments retain a biological activity of the receptor, including antigenic activity. Variants of the defined sequence and fragments also form part of the present invention. Preferred variants are those that vary from the referents by conservative amino acid substitutions -- i.e., those that substitute a residue with another of like characteristics. Typical such substitutions are among Ala. Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln; and among the basic residues Lys and Arg; or aromatic residues Phe and Tyr. Particularly preferred are variants in which several, 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination.

The receptor polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

Polynucleotides of the Invention

Another aspect of the invention relates to receptor polynucleotides. Receptor polynucleotides include isolated polynucleotides which encode the receptor polypeptides and fragments, and polynucleotides closely related thereto. More specifically, a receptor polynucleotide of the invention includes a polynucleotide comprising the nucleotide sequence contained in SEQ ID NO:X encoding a receptor polypeptide of SEQ ID NO:Y, and polynucleotide having the particular sequence of SEQ ID NO:X.

Receptor polynucleotides further include a polynucleotide comprising a nucleotide sequence that has at least 80% identity over its entire length to a nucleotide sequence encoding the receptor polypeptide of SEQ ID NO:Y, and a polynucleotide comprising a nucleotide sequence that is at least 80% identical to that of SEQ ID NO:X over its entire length. In this regard, polynucleotides at least 90% identical are particularly preferred, and those with at least 95% are especially preferred. Furthermore, those with at least 97% are highly preferred and those with at least 98-99% are most highly preferred, with at least 99% being the most preferred. Also included under receptor polynucleotides are a nucleotide sequence which has sufficient identity to a nucleotide sequence contained in SEQ ID NO:X, or contained in the cDNA insert in the plasmid deposited with ATCC, to hybridize under conditions useable for amplification or for use as a probe or marker. Moreover, the receptor polynucleotide includes a nucleotide sequence having at least 80% identity to a nucleotide sequence encoding the receptor polypeptide expressed by the cDNA insert deposited at the ATCC, and a nucleotide sequence comprising at least 15 contiguous nucleotides of such cDNA insert. In this regard, polynucleotides at least 90% identical are particularly preferred, and those with at least 95% are especially preferred. Furthermore, those with at least 97% are highly preferred and those with at least 98-99% are most highly preferred, with at least 99% being the most preferred. The invention also provides polynucleotides which are complementary to all the above receptor polynucleotides.

The receptors of the invention are structurally related to other proteins of specified receptor families, as shown by the results in the Figures. The cDNA sequence of SEQ ID NO:X encodes a polypeptide as described in Table 1 as SEQ ID NO:Y. Because the receptor polypeptides contain domains similar in structure to other receptor family members, the receptors of the present invention are expected to have,

inter alia, similar biological functions/properties to their homologous polypeptides and polynucleotides, and their utility is obvious to anyone skilled in the art.

Table 1

Clone ID Name	SEQ ID NO:X	SEQ ID NO:Y	ATCC Deposit No.	ATCC Deposit Date	Receptor Family	Homology
HMACR70	1	18	209054 ##### 01/21/98	05/16/97	Ig	Sialoadhesin OB-1
HTEDK48			209054	05/16/97	TM4SF	MRC-OX44 PETA-3
1-1849 bp	2					
160-900 bp	3	19				
HTPED39 HPWAE25	4	20	209054 #####	05/16/97 1/21/98	TM4SF	NAG-2 TALLA-1
HTPEF86	5	21	209053	05/16/97	TM4SF	CD20 B1 Antigen
HSBBF02	6	22	209054	05/16/97	TM4SF	TALLA-1
HLTAH80	7	23	97242 209054	08/02/95 05/16/97	TM4SF	TALLA-1
HTPBA27	8	24	97242 209054	08/02/95 05/16/97	TM4SF	NAG-2
HAIDQ59			209054	05/16/97	TM4SF	CD9 Antigen
5' Sequence	9	25				
3' Sequence	10					
HHFEK40	11	26	209054	05/16/97	TM4SF	PETA-3
HGBGV89	12	27	209125 209054	06/09/97 05/16/97	TM4SF	L6H
HUVBB80	13	28	209054	05/16/97	TM4SF	L6
HJACE54	14	29	209053	05/16/97	Lectin	Galectin-3 Galectin-5 Galectin-8
HROAD63	15	30	209053	05/16/97	Ly6	E48 splice variant
HMWGS46	16	31	209053	05/16/97	Prohibitin	BAP-37
HNFGW06	17	32	209053	05/16/97	EGFR	EGFR

The novel full-length cDNA clone designated **HMACR70** may be a member of the sialoadhesin family of the Ig superfamily of receptor-like molecules and a CD33 homologue. HMACR70 contains a 1497 nucleotide cDNA insert encoding a 315 amino acid ORF and was cloned from a GM-CSF-treated human macrophage cDNA library. The only additional cDNA libraries in the HGS database which include this clone are human eosinophils and possibly human gall bladder. A BLAST analysis of the amino acid sequence of HMACR70 demonstrates that this clone exhibits approximately 50% identity and 69% similarity over a 300 amino acids stretch of a gene termed human

differentiation antigen, and 38% identity and 62% similarity of the human myelin-associated glycoprotein precursor CD33 gene.

A more recent BLAST analysis confirms HMACR70's designation as a sialoadhesin family member. HMACR70 is homologous to two recently identified sialoadhesin family members, human OB binding protein (OB) 1 and 2. (See, Genbank Accession No. U71382; see Figure 1.) It is thought that OB-1 and OB-2 may bind leptin. Thus, HMACR70, as a sialoadhesin family member, may act to attenuate or even amplify intercellular routes of communication, including binding to leptin or modulating the activity of immune cells, such as macrophages. Clearly, any diseases affected by these processes could be treated by the polypeptide or fragment of HMACR70.

The full-length nucleotide sequences of ten novel human cDNA clones which potentially belong to the TM4SF superfamily are disclosed in the table above and will be addressed sequentially.

The cDNA clone **HTEDK48** contains a 1849 nucleotide cDNA insert encoding a 245 amino acid ORF that was cloned from a human testes cDNA library. The coding sequence of HTEDK48 (SEQ ID NO: 3) may be fused to other human proteins, such as 3-hydroxyacyl-CoA dehydrogenase. BLAST analysis of the amino acid sequence of HTEDK48 demonstrates that this clone exhibits approximately 30% identity and 51% similarity over a 245 amino acid stretch of the CD82 molecule. Recent studies have shown that CD82 can associate with CD4 or CD8 and deliver costimulatory signals for the TCR/CD3 pathway. CD82 has also been found to be involved in syncytium formation in HTLV-I-infected T-cells. And finally, in a recently published study in which the expression of the CD82 gene by tumors of the lung was examined retrospectively, it was reported that CD82 may be linked to the suppression of tumor metastasis of prostate cancer. The study also reported that decreased CD82 expression may be involved in malignant progression of such cancers. Thus, HTEDK48 may also be involved in the development of cancer.

A more recent BLAST analysis shows that HTEDK48 is homologous the rat leukocyte antigen, MRC OX-44, and the platelet endothelial tetraspan antigen -3 (PETA-3). (See Figure 2X.) MRC OX-44, a member of a new family of cell surface proteins, appears to be involved in growth regulation. (See, Bellacosa, A., et al., "The Rat Leukocyte antigen MRC OX-44 is a Member of a New Family of Cell Surface Proteins which Appear to be Involved in Growth Regulation," Mol. Cell. Bio. 11: 2864-2872 (1991).) Similarly, PETA-3 has been located to platelet endothelial cells, and an anti-PETA-3 antigen monoclonal antibody can stimulate platelet aggregation and mediator release. (See, Fitter, S., "Molecular Cloning of cDNA Encoding a Novel Platelet-Endothelial Cell Tetra-Span Antigen, PETA-3," Blood, 86(4):1348-1355 (1995).) Thus, HTEDK48 may function similar to MRC OX-44 or PETA-3 to affect

growth of blood cells. Administering polypeptides or fragments of HTEDK48 may be an effective treatment of blood disorders.

The cDNA clone **HPWAE25** contains a 1288 nucleotide cDNA insert encoding a 273 amino acid ORF that was cloned from a human pancreas tumor cDNA library, while clone **HTPED39** represents a truncated cDNA sequence. This clone also appears in a number of other cDNA libraries constructed from a variety of human cell and tissue types including keratinocytes, ulcerative colitis, striatum depression, lymph node breast cancer, ovarian cancer, stage B2 prostate cancer, kidney medulla, and others. Northern blot analysis of HLTAH80 also shows expression in a variety of human cell lines including U937, MM96, WM115, and MDAMB231. A BLAST analysis of the amino acid sequence of HTPED39 demonstrates that this clone exhibits approximately 35% identity and 50% similarity over the entire length of the CD37 molecule. The CD37 antigen is expressed on B cells and on a subpopulation of T cells, but not on pre-B or plasma cells. It has been reported that CD37 expression is downregulated in conjunction with B-cell activation, suggesting that CD37 may be involved in the processes which dictate the activation state of the B-cell.

Moreover, HPWAE25 is also homologous to recently identified TM4SF members, NAG-2 and TALLA-1. (See Figure 3.) NAG-2 is thought to complex with integrins and other TM4SF proteins, while TALLA-1 is a highly specific marker of T-cell acute lymphoblastic leukemia and neuroblastoma. (See, Tachibana, I., et al., "NAG-2, A Novel Transmembrane-4 Superfamily (TM4SF) Protein that Complexs with Integrins and Other TM4SF Proteins," J. Biol. Chem., 272:29181-29189 (1997); Takagi, S., "Identification of a Higly Specific Surface Marker of T-cell Acute Lymphoblastic Leukemia and Neuroblastoma as a New Member of the Transmembrane 4 Superfamily," Int. J. Cancer 61(5):706-715 (1995).) Thus, HPWAE25 may be involved the development of cancer, particularly leukemia, lymphoma, and neuroblastoma. HPWAE25 may be used as an effective treatment of these cancers, as well as a diagnostic marker.

A subfamily of TM4SF receptors include CD20 proteins. A CD20-like cDNA clone was obtained from a human pancreas tumor cDNA library and contains a 1236 nucleotide insert which encodes a 250 amino acid ORF. A BLAST analysis of the deduced amino acid sequence of **HTPEF86** exhibits approximately 41% identity and 61% similarity to the CD20 gene, also known as B1 antigen. (See Figure 4.) Expression of this gene is detected in only two additional HGS human cDNA libraries; amygdala depression and 9 week early stage human. Although the precise functional role of CD20 has yet to be determined, it is clear that CD20 plays a key role in the regulation of B-cell activation. Based primarily on sequence identity, the novel CD20-like molecule presented herein may also be involved in cell cycle activation. Potential therapeutic and/or diagnostic applications for HTPEF86 may include such clinical

presentations as juvenile rheumatoid arthritis, Graves' Disease, and a number of B-cell lymphomas or other lymphoid tumors.

The clone **HSBBF02** contains a 1115 nucleotide cDNA insert encoding a 245 amino acid ORF and was cloned from an HSC 172 cell line cDNA library. This clone also appears in a number of other cDNA libraries constructed from a variety of human cell and tissue types including brain amygdala depression, endothelial cells, fetal liver and heart, osteoblasts, testes, and others. A BLAST analysis of the amino acid sequence of HSBBF02 demonstrates that this clone exhibits approximately 64% identity and 80% similarity with the A15 molecule over a 131 amino acid stretch (A15 is composed of 244 amino acids). A more recent BLAST search shows that HSBBF02 is similar to the TALLA-1 protein and may in fact be a closely related family member. (See Figure 5.)

In addition, a second cDNA clone, designated **HLTAH80**, exhibits sequence similarity to the A15 molecule and TALLA-1. (See Figure 6.) This clone contains a 1662 nucleotide cDNA insert encoding a 253 amino acid ORF and was cloned from a human T-cell lymphoma cDNA library. This clone also appears in a number of other cDNA libraries constructed from a variety of human cell and tissue types including B-cell lymphoma, corpus collosum, endometrial tumor, osteosarcoma, testes, and others. Northern blot analysis of HLTAH80 also shows expression in a variety of human tissues including spleen, lymph node, thymus, PBLs, heart, and a particularly strong signal in skeletal muscle and pancreas. A BLAST analysis of the amino acid sequence of HLTAH80 demonstrates that this clone exhibits approximately 35% identity and 55% similarity over the entire length of the A15 molecule.

Since expression of A15 drops to undetectable levels when comparing immature T-cells to peripheral blood lymphocytes, it is thought that A15 may play a role in the development of T-cells. Furthermore, the MXS1(CCG-B7) gene which codes for A15 contains a number of triplet nucleotide repeats which have been associated with neuropsychiatric diseases such as Huntington's chorea, fragile X syndrome, and myotonic dystrophy. In addition, A15 appears to be expressed exclusively on T-cell acute lymphoblastic leukemia cell lines, including several derived from adult T-cell leukemia and those established by immortalization with human T-cell leukemia virus type 1 or Herpesvirus saimiri. Thus, clones HLTAH80 and/or HSBBF02 may also be involved in diseases caused by the expansion of repeats or chromosomal instability.

The cDNA clone **HTPBA27** contains a 1345 nucleotide cDNA insert encoding a 238 amino acid ORF and was cloned from a human tumor pancreas cDNA library. This clone also appears in a number of other cDNA libraries constructed from a variety of human cell and tissue types including cerebellum, breast lymph node, osteosarcoma, adult testes, RS4;11 bone marrow cell line, microvascular endothelial cells, and others. A BLAST analysis of the amino acid sequence of HTPBA27 demonstrates that this

clone exhibits approximately 40% identity and 64% similarity with a glycoprotein termed CD53 over its entire length. CD53 is thought to be involved in thymopoiesis, since rat CD53 can be detected on immature CD4-8-thymocytes and the functionally mature single-positive subset, but cannot be detected on the intermediate CD4+8+ thymocytic subset of cells. The CD53 molecule has also been implicated as a component of signal transduction pathways in B cells, monocytes and granulocytes, rat macrophages, NK, and T cells. Moreover, as illustrated in Figure 7, HTPBA27 was recently confirmed as a TM4SF receptor. (See, Tachibana, I., et al., "NAG-2, A Novel Transmembrane-4 Superfamily (TM4SF) Protein that with Integrins and Other TM4SF Proteins," J. Biol. Chem., 272:29181-29189 (1997).) Calling the HTPBA27 polypeptide NAG-2, this group confirmed HTPBA27's status as a TM4SF receptor by showing that NAG-2 complexes with integrin and other TM4SF receptors. Thus, diseases caused by the failure of HTPBA27 to complex with integrin and other TM4SF receptors can be treated by administering HTPBA27. HTPBA27 can also be used to diagnose these diseases.

The cDNA clone **HAIDQ59** contains cDNA insert encoding a 221 amino acid ORF that was cloned from a human epithelial cell induced with TNFa and INF cDNA library. The 5' end of HAIDQ59 is represented by the SEQ ID NO: 9, while the 3' end is represented by SEQ ID NO: 10. This clone appears in only two additional cDNA libraries in the HGS database. These two libraries were constructed from the human Jurkat T-cell line and human microvascular endothelial cells. A BLAST analysis of the amino acid sequence of HAIDQ59 demonstrates that this clone exhibits approximately 53% identity and 69% similarity over 226 amino acids of the CD9 TM4SF molecule. (See Figure 8.) It has been demonstrated that the CD9 molecule is involved in signal transduction pathways in platelets, as well as in cell adhesion in both platelets and pre-B-cell lines. Intriguingly, a monoclonal antibody (vpg15), which recognizes the feline homologue of CD9, has been shown to block infection by feline immunodeficiency virus (FIV). Furthermore, a recent study shows that cells expressing high levels of CD9 exhibited suppressed cell motility. Thus, HAIDQ59 may also be involved in signal transduction of blood cells.

The cDNA clone **HHFEK40** contains a 936 nucleotide cDNA insert encoding a 252 amino acid ORF and was cloned from a human fetal heart cDNA library. This clone appears once in the human fetal heart cDNA library and possibly in a hemangiopericytoma cDNA library. A BLAST analysis of the amino acid sequence of HHFEK40 demonstrated that this clone exhibits approximately 60% identity and 75% similarity over the entire length of a molecule designated PETA-3. (See Figure 9.) PETA-3 was originally identified as a novel human platelet surface glycoprotein termed gp27. Although PETA-3 is present in low abundance on the platelet surface, an anti-PETA-3 monoclonal antibody can stimulate platelet aggregation and mediator release.

Thus, HHFEK40 may function similar to PETA-3 to affect growth of blood cells. Administering polypeptides or fragments of HHFEK40 may be an effective treatment of blood disorders.

The cDNA clone **HGBGV89** contains a 738 nucleotide cDNA insert encoding a 197 amino acid ORF and was cloned from a human gall bladder cDNA library. The only two additional appearances of this clone in the HGS database are in a normalized fetal liver cDNA library and in a fetal liver/spleen cDNA library. The cDNA clone **HUVBB80** contains a 1071 nucleotide cDNA insert encoding a 201 amino acid ORF and was cloned from a human umbilical vein cDNA library. This clone appears in several additional cDNA libraries in the HGS database including prostate BPH, thyroid, and fetal liver/spleen. BLAST analyses of the amino acid sequences of HGBGV89 and HUVBB80 demonstrate that these clones exhibit approximately 49% identity and 65% similarity and 47% identity and 68% similarity, respectively, over the entire length of a molecule designated L6 surface protein or human tumor-associated antigen L6. (See Figures 10 & 11.) Moreover, another group has confirmed the TM4SF receptor homology of HGBGV89 by describing the protein as a putative transmembrane protein L6H. (See Genbank Accession No 2587054; see Figure 10.) The L6 cell surface antigen is highly expressed on lung, breast, colon, and ovarian carcinomas. Promising results of phase 1 clinical studies have been reported with an anti-L6 monoclonal antibody, or its humanized counterpart, suggesting that the L6 antigen may be an attractive target for monoclonal antibody-based cancer therapy.

In summary, there is a clear need for identifying and exploiting novel members of the TM4SF superfamily such as those described herein. Although structurally related, these factors will likely possess diverse and multifaceted functions in a variety of cell and tissue types. Receptor type molecules, such as the novel potential members of the TM4SF superfamily detailed here, should prove useful in target based screens for small molecules and other such pharmacologically valuable factors. Monoclonal antibodies raised against such factors may prove useful as therapeutics in an anti-tumor, diagnostic, or other capacity. Furthermore, factors such as the nine novel TM4SF superfamily-like molecules described here may prove useful in an active or passive immunotherapeutical role in patients with cancer or other immunocompromised disease states.

Besides TM4SF receptors, receptors from other families are also described. For example, clone **HJACE54**, also called galectin 11, exhibits significant sequence identity to the rat galectin 5, the chicken galectin 3 gene, and the human galectin 8 genes. (See Figure 12.) The galectin 11 cDNA clone contains an 865 nucleotide insert which encodes a 133 amino acid ORF. The clone was obtained from a Jurkat T-cell G1 phase cDNA library. A BLAST analysis of the deduced amino acid sequence of HJACE54 demonstrates approximately 35% identity and 57% similarity to the amino

acid sequence of the rat galectin 5 gene. Expression of galectin 11 is quite limited in the HGS database. In fact, the only two additional ESTs in the HGS database which contain the HJACE54 sequence were found in human neutrophil and human infant adrenal gland cDNA libraries. Northern blot analyses have not been performed to examine expression patterns of the galectin 11 gene.

Various galectins have been shown to function in the mechanisms of intercellular communication. For example, depending on cell type, galectin 1 has been observed to modulate cell adhesion either positively or negatively. More specifically, galectin 1 appears to inhibit cell adhesion of skeletal muscle presumably by galectin 1-mediated disruption of laminin-integrin a_7b_1 interactions. Alternatively, galectin 1 appears to promote cell adhesion in several non-skeletal muscle cell types examined presumably by a glycoconjugate cross-linking mechanism. Galectin 3 has also been observed to function in modulating cell-adhesion, as well as in the activation of certain immune cells by cross-linking IgE and IgE receptors. In addition, galectins have been observed to be involved in the regulation of immune cell activity, as well as in such diverse processes as cell adhesion, proliferation, inflammation, autoimmunity, and metastasis of tumor cells. Furthermore, a galectin-like antigen designated HOM-HD-21 was recently found to be highly expressed in a Hodgkin's Disease cDNA library. Very recently, a novel galectin, termed PCTA-1, was identified as a specific cell surface marker on human prostate cancer cell lines and patient-derived carcinomas. Galectins have also been found to function intracellularly as a component of ribonucleoprotein complexes. Finally, galectins 1 and 3 have each been found to modulate T-cell growth and apoptosis by interaction with CD45 and possibly Bcl2, respectively. As a result, the discovery of a novel galectin, such as that encoded by HJACE54, is likely to be a valuable asset both diagnostically and therapeutically.

Additionally, a full-length nucleotide sequence of a novel human cDNA clone which encodes an apparent splice variant of the previously described human E48 antigen has recently been determined. (See Figure 13.) Clone **HROAD63** contains a 441 nucleotide cDNA which encodes a 70 amino acid polypeptide. This novel clone exhibits significant sequence identity to several members of a relatively new family of cell-surface proteins termed the Ly6 superfamily. These members include murine and human SCA-2, rat Ly-6 (also termed ThB), and human CD59 [also known as protectin or membrane attack complex inhibition factor (MACIF)]. The novel E48 splice variant was obtained from the HGS human stomach cDNA library. The clone is present in only a limited number of other HGS cDNA libraries including kidney cancer, keratinocyte, and tongue. An alignment of the nucleotide sequences of the human E48 and HROAD63 cDNAs demonstrates that the initial 168 and 178 nucleotides of E48 and HROAD63, respectively, are identical, with the exception of an additional 10 nucleotides of sequence at the extreme 5' end of the HROAD63 sequence. The

sequence of the two clones is also identical for an additional 229 nucleotides including the 3' end of the coding sequences and the entire 3' untranslated regions. The only divergence of nucleotide sequence in this region of the clones is the deletion of a single thymidine residue in the 3' UTR of the E48 cDNA. The major difference between the two nucleotide sequences is a 329 nucleotide deletion from the HROAD63 sequence. This deletion causes a shift in the HROAD63 reading frame and encompasses the translational stop signal used in the E48 clone. As a result, the carboxy terminal sequence of HROAD63 is radically altered with regard to that of E48 (as illustrated in Figure 13 by the obvious differences between amino acids 56-128 of E48 and 56-70 of HROAD63 in the amino acid alignment). The clinical presentation of disorders, including abnormal skin and hair phenotypes, may be attributed, at least in part, to a non-functional Ly6 superfamily member such as E48 or HROAD63. HROAD63 may also be involved in blood disorders, as seen with its homologues SCA-2 and CD59.

A novel prohibitin cDNA clone presented herein was originally identified in a human bone marrow cell line (RS4;11) cDNA library. The clone contains a 1066 nucleotide insert which encodes a 299 amino acid polypeptide. BLAST and BestFit analyses of the predicted amino acid sequence of **HMWGS46** demonstrate a highly significant sequence identity to a murine protein termed IgM B-cell receptor associated protein (BAP)-37 (Genbank accession number X78683). The HMWGS46 amino acid sequence exhibits nearly perfect identity and similarity over the entire length of the murine BAP-37 sequence. (See Figure 14.) In addition, the full-length nucleotide sequences of HMWGS46 and BAP-37 exhibit at least 87% identical. The HMWGS46 clone also exhibits approximately 49% sequence identity and 85% sequence similarity to a human gene designated prohibitin. Finally, the HMWGS46 cDNA appears in a substantial number of HGS human cDNA libraries in addition to the bone marrow cell line cDNA library from which it was cloned. Some of the cDNA libraries in which this clone appears include keratinocytes, induced endothelial cells, activated neutrophils, synovial sarcoma, colon carcinoma cell line, Jurkat cell line membrane bound polysomes, epileptic frontal cortex, primary dendritic cells, and a number of others. The novel gene related to prohibitin and BAP-37 may prove quite useful as a diagnostic for tumorigenesis, as well as a target for therapeutic intervention of such an event. Thus, although the precise functional role of the prohibitin family members are less than clear, it is quite likely that such homologues are involved in such complex processes as development, senescence, and tumor suppression. Therefore a novel gene, such as HMWGS46, may prove quite useful as a diagnostic for tumorigenesis, as well as a target for therapeutic intervention of such an event.

A human cDNA clone encoding a novel epidermal growth factor receptor (EGFR)-like molecule is also disclosed. The novel EGFR-like cDNA clone presented herein was originally identified in an activated human neutrophil cDNA library. The

clone contains a 704 nucleotide insert which encodes a 168 amino acid polypeptide. A BLAST analysis of the predicted amino acid sequence of **HNFGW06** demonstrates that this novel clone exhibits approximately 85% identity and 90% similarity to a protein designated epidermal growth factor receptor-related protein [Homo sapiens]. (See Figure 15.) The expression profile of the HNFGW06 clone in the HGS database indicates the existence of a fairly highly restricted expression pattern. In addition to the activated neutrophil library from which this clone was obtained, it also appears in the following HGS human cDNA libraries: synovial sarcoma, smooth muscle, placenta, and possibly primary dendritic cells.

The novel EGFR-like cDNA clone HNFGW06 may lead to a number of exciting possibilities for therapeutic and/or diagnostic treatments or reagents. For example, HNFGW06 may be involved in the onset of human breast cancers as well. In addition, due to the fact that TGF-a acts through binding to the EGFR, it is possible that HNFGW06 may also play a role in a variety of gastric processes including regulation of acid secretion, regulation of mucous cell growth, and protection against ethanol- and aspirin-induced injury to gastric tissues.

GENERATING POLYNUCLEOTIDES

Polynucleotides of the present invention encoding a receptor may be obtained using standard cloning and screening, from a cDNA library derived from mRNA in cells specified in Table 1 using the expressed sequence tag (EST) analysis (Adams, M.D., et al. Science (1991) 252:1651-1656; Adams, M.D. et al., Nature, (1992) 355:632-634; Adams, M.D., et al., Nature (1995) 377 Supp:3-174.) Polynucleotides of the invention can also be obtained from natural sources such as genomic DNA libraries or can be synthesized using well known and commercially available techniques.

The nucleotide sequence encoding a receptor polypeptide of SEQ ID NO:Y may be identical to the polynucleotide encoding SEQ ID NO:Y, or it may be a sequence, which as a result of the redundancy (degeneracy) of the genetic code, also encodes the polypeptide of SEQ ID NO:Y.

When the polynucleotides of the invention are used for the recombinant production of a receptor polypeptide, the polynucleotide may include the coding sequence for the mature polypeptide or a fragment thereof, by itself; the coding sequence for the mature polypeptide or fragment in reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence, or other fusion peptide portions. For example, a marker sequence which facilitates purification of the fused polypeptide can be encoded. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in

Gentz et al., Proc Natl Acad Sci USA (1989) 86:821-824, or is an HA tag. The polynucleotide may also contain non-coding 5' and 3' sequences, such as transcribed, non-translated sequences, splicing and polyadenylation signals, ribosome binding sites and sequences that stabilize mRNA.

Further preferred embodiments are polynucleotides encoding receptor variants comprising the amino acid sequence of receptor polypeptide of Table 1 (SEQ ID NO:Y) in which several, 5-10, 1-5, 1-3, 1-2 or 1 amino acid residues are substituted, deleted or added, in any combination.

The present invention further relates to polynucleotides that hybridize to the herein above-described sequences. In this regard, the present invention especially relates to polynucleotides which hybridize under stringent conditions to the herein above-described polynucleotides. As herein used, the term "stringent conditions" means hybridization will occur only if there is at least 80%, and preferably at least 90%, and more preferably at least 95%, yet even more preferably 97-99% identity between the sequences.

Polynucleotides of the invention, which are identical or sufficiently identical to a nucleotide sequence contained in SEQ ID NO:X or a fragment thereof, or to the cDNA insert in the plasmid deposited at the ATCC, or a fragment thereof, may be used as hybridization probes for cDNA and genomic DNA, to isolate full-length cDNAs and genomic clones encoding the receptor and to isolate cDNA and genomic clones of other genes (including genes encoding homologs and orthologs) that have a high sequence similarity to the receptor gene. Such hybridization techniques are known to those of skill in the art. Typically these nucleotide sequences are 80% identical, preferably 90% identical, more preferably 95% identical to that of the referent. The probes generally will comprise at least 15 nucleotides. Preferably, such probes will have at least 30 nucleotides and may have at least 50 nucleotides. Particularly preferred probes will range between 30 and 50 nucleotides.

In one embodiment, to obtain a polynucleotide encoding the receptor polypeptide, including homologs and orthologs from other species, comprises the steps of screening an appropriate library under stringent hybridization conditions with a labeled probe having the SEQ ID NO:X or a fragment thereof; and isolating full-length cDNA and genomic clones containing said polynucleotide sequence. Such hybridization techniques are well known to those of skill in the art. Stringent hybridization conditions are as defined above or, alternatively, conditions under overnight incubation at 42°C in a solution comprising: 50% formamide, 5xSSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10 % dextran sulfate, and 20 microgram/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

The polynucleotides and polypeptides of the present invention may be employed as research reagents and materials for discovery of treatments and diagnostics to animal and human disease.

Vectors, Host Cells, Expression

The present invention also relates to vectors which comprise a polynucleotide or polynucleotides of the present invention, and host cells which are genetically engineered with vectors of the invention and to the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention.

For recombinant production, host cells can be genetically engineered to incorporate expression systems or portions thereof for polynucleotides of the present invention. Introduction of polynucleotides into host cells can be effected by methods described in many standard laboratory manuals, such as Davis et al., BASIC METHODS IN MOLECULAR BIOLOGY (1986) and Sambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989) such as calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction or infection.

Representative examples of appropriate hosts include bacterial cells, such as streptococci, staphylococci, E. coli, Streptomyces and Bacillus subtilis cells; fungal cells, such as yeast cells and Aspergillus cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, HEK 293 and Bowes melanoma cells; and plant cells.

A great variety of expression systems can be used. Such systems include, among others, chromosomal, episomal and virus-derived systems, e.g., vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression systems may contain control regions that regulate as well as engender expression. Generally, any system or vector suitable to maintain, propagate or express polynucleotides to produce a polypeptide in a host may be used. The appropriate nucleotide sequence may be inserted into an expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook et al., MOLECULAR CLONING, A LABORATORY MANUAL (supra).

For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the desired polypeptide. These signals may be endogenous to the polypeptide or they may be heterologous signals.

If the receptor polypeptide is to be expressed for use in screening assays, generally, it is preferred that the polypeptide be produced at the surface of the cell. In this event, the cells may be harvested prior to use in the screening assay. If the receptor polypeptide is secreted into the medium, the medium can be recovered in order to recover and purify the polypeptide; if produced intracellularly, the cells must first be lysed before the polypeptide is recovered.

Receptor polypeptides can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during isolation and or purification.

Diagnostic Assays

This invention also relates to the use of receptor polynucleotides or polypeptides for use as diagnostic reagents. Detection of a mutated form of the receptor gene associated with a dysfunction will provide a diagnostic tool that can add to or define a diagnosis of a disease or susceptibility to a disease which results from under-expression, over-expression or altered expression of the receptor. Individuals carrying mutations in the receptor gene may be detected at the DNA level by a variety of techniques.

Nucleic acids for diagnosis may be obtained from a subject's cells, such as from blood, urine, saliva, tissue biopsy or autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR or other amplification techniques prior to analysis. RNA or cDNA may also be used in similar fashion. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to labeled receptor nucleotide sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence differences may also be detected by alterations in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing. (See, e.g., Myers et al., Science (1985) 230:1242.) Sequence changes at specific locations may also be revealed by

nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method. (See Cotton et al., Proc Natl Acad Sci USA (1985) 85: 4397-4401.) In another embodiment, an array of oligonucleotides probes comprising receptor nucleotide sequence or fragments thereof can be constructed to conduct efficient screening of e.g., genetic mutations. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability. (See for example: M.Chee et al., Science, Vol 274, pp 610-613 (1996).)

The diagnostic assays offer a process for diagnosing or determining a susceptibility to specific diseases through detection of mutation in the receptor gene by the methods described.

In addition, specific diseases can be diagnosed by methods comprising determining from a sample derived from a subject an abnormally decreased or increased level of receptor polypeptide or receptor mRNA. Decreased or increased expression can be measured at the RNA level using any of the methods well known in the art for the quantitation of polynucleotides, such as, for example, PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods. Assay techniques that can be used to determine levels of a protein in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays.

Thus in another aspect, the present invention relates to a diagnostic kit for a disease or susceptibility to a disease which comprises:

(a) a receptor polynucleotide, preferably the nucleotide sequence of SEQ ID NO:X, or a fragment thereof ;

(b) a nucleotide sequence complementary to that of (a);

(c) a receptor polypeptide, preferably the polypeptide of SEQ ID NO:Y, or a fragment thereof; or

(d) an antibody to a receptor polypeptide, preferably to the polypeptide of SEQ ID NO: Y.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component.

Chromosome Assays

The nucleotide sequences of the present invention are also valuable for chromosome identification. The sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes according to the present invention is an important first step in correlating those sequences with gene associated disease. Once a sequence has been mapped to a precise chromosomal location, the physical position of the

sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

The differences in the cDNA or genomic sequence between affected and unaffected individuals can also be determined. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent of the disease.

Antibodies

The polypeptides of the invention or their fragments or analogs thereof, or cells expressing them can also be used as immunogens to produce antibodies immunospecific for the receptor polypeptides. The term "immunospecific" means that the antibodies have substantially greater affinity for the polypeptides of the invention than their affinity for other related polypeptides in the prior art.

Antibodies generated against the receptor polypeptides can be obtained by administering the polypeptides or epitope-bearing fragments, analogs or cells to an animal, preferably a nonhuman, using routine protocols. For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler, G. and Milstein, C., Nature (1975) 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor et al., Immunology Today (1983) 4:72) and the EBV-hybridoma technique (Cole et al., MONOCLONAL ANTIBODIES AND CANCER THERAPY, pp. 77-96, Alan R. Liss, Inc., 1985).

Techniques for the production of single chain antibodies (U.S. Patent No. 4,946,778) can also be adapted to produce single chain antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms including other mammals, may be used to express humanized antibodies.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or to purify the polypeptides by affinity chromatography. Antibodies against receptor polypeptides may also be employed to treat diseases.

Vaccines

Another aspect of the invention relates to a method for inducing an immunological response in a mammal which comprises inoculating the mammal with a receptor polypeptide, or a fragment thereof, adequate to produce antibody and/or T cell immune response to protect said animal from a disease. Yet another aspect of the invention relates to a method of inducing immunological response in a mammal which

comprises, delivering a receptor polypeptide via a vector directing expression of the receptor polynucleotide in vivo in order to induce such an immunological response to produce antibody to protect said animal from diseases.

Further aspect of the invention relates to an immunological/vaccine formulation (composition) which, when introduced into a mammalian host, induces an immunological response in that mammal to a receptor polypeptide wherein the composition comprises a receptor polypeptide or receptor gene. The vaccine formulation may further comprise a suitable carrier. Since a receptor polypeptide may be broken down in the stomach, it is preferably administered parenterally (including subcutaneous, intramuscular, intravenous, intradermal etc. injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation instonic with the blood of the recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

Screening Assays

The receptor polypeptide of the present invention may be employed in a screening process for compounds which bind the receptor and which activate (agonists) or inhibit activation of (antagonists) the receptor polypeptide of the present invention. Thus, polypeptides of the invention may also be used to assess the binding of small molecule substrates and ligands in, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These substrates and ligands may be natural substrates and ligands or may be structural or functional mimetics. See Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).

The receptor polypeptides are responsible for many biological functions, including many pathologies. Accordingly, it is desirous to find compounds and drugs which stimulate the receptor on the one hand and which can inhibit the function of the receptor on the other hand. In general, agonists are employed for therapeutic and prophylactic purposes for such conditions and diseases. Antagonists may be employed for a variety of therapeutic and prophylactic purposes for such conditions and diseases.

In general, such screening procedures involve producing appropriate cells which express the receptor polypeptide of the present invention on the surface thereof.

Such cells include cells from mammals, yeast, Drosophila or E. coli. Cells expressing the receptor (or cell membrane containing the expressed receptor) are then contacted with a test compound to observe binding, or stimulation or inhibition of a functional response.

The assays may simply test binding of a candidate compound wherein adherence to the cells bearing the receptor is detected by means of a label directly or indirectly associated with the candidate compound or in an assay involving competition with a labeled competitor. Further, these assays may test whether the candidate compound results in a signal generated by activation of the receptor, using detection systems appropriate to the cells bearing the receptor at their surfaces. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist by the presence of the candidate compound is observed.

Further, the assays may simply comprise the steps of mixing a candidate compound with a solution containing a receptor polypeptide to form a mixture, measuring receptor activity in the mixture, and comparing the receptor activity of the mixture to a standard.

The receptor cDNA, protein and antibodies to the protein may also be used to configure assays for detecting the effect of added compounds on the production of receptor mRNA and protein in cells. For example, an ELISA may be constructed for measuring secreted or cell associated levels of receptor protein using monoclonal and polyclonal antibodies by standard methods known in the art, and this can be used to discover agents which may inhibit or enhance the production of the receptor (also called antagonist or agonist, respectively) from suitably manipulated cells or tissues. Standard methods for conducting screening assays are well understood in the art.

Examples of potential receptor antagonists include antibodies or, in some cases, oligonucleotides or proteins which are closely related to the ligand of the receptor, e.g., a fragment of the ligand, or small molecules which bind to the receptor but do not elicit a response, so that the activity of the receptor is prevented.

Thus in another aspect, the present invention relates to a screening kit for identifying agonists, antagonists, ligands, receptors, substrates, enzymes, etc. for receptor polypeptides; or compounds which decrease or enhance the production of receptor, which comprises:

(a) a receptor polypeptide, preferably that of SEQ ID NO:Y;

(b) a recombinant cell expressing a receptor polypeptide, preferably that of SEQ ID NO:Y;

(c) a cell membrane expressing a receptor polypeptide; preferably that of SEQ ID NO: Y; or

(d) antibody to a receptor polypeptide, preferably that of SEQ ID NO: Y.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component.

Prophylactic and Therapeutic Methods

This invention provides methods of treating an abnormal conditions related to both an excess of and insufficient amounts of receptor activity.

If the activity of the receptor is in excess, several approaches are available. One approach comprises administering to a subject an inhibitor compound (antagonist) as described along with a pharmaceutically acceptable carrier in an amount effective to inhibit activation by blocking the binding of ligands to the receptor or by inhibiting a second signal, and thereby alleviating the abnormal condition.

In another approach, soluble forms of the receptor polypeptides still capable of binding the ligand in competition with endogenous receptor may be administered. Typical embodiments of such competitors comprise fragments of the receptor polypeptide.

In still another approach, expression of the gene encoding endogenous receptor can be inhibited using expression blocking techniques. Known such techniques involve the use of antisense sequences, either internally generated or separately administered. (See, for example, O'Connor, J Neurochem (1991) 56:560 in Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Alternatively, oligonucleotides which form triple helices with the gene can be supplied. (See, for example, Lee et al., Nucleic Acids Res (1979) 6:3073; Cooney et al., Science (1988) 241:456; Dervan et al., Science (1991) 251:1360.) These oligomers can be administered per se or the relevant oligomers can be expressed in vivo.

For treating abnormal conditions related to an under-expression of the receptor and its activity, several approaches are also available. One approach comprises administering to a subject a therapeutically effective amount of a compound which activates the receptor, i.e., an agonist as described above, in combination with a pharmaceutically acceptable carrier, to thereby alleviate the abnormal condition. Alternatively, gene therapy may be employed to effect the endogenous production of the receptor by the relevant cells in the subject. For example, a polynucleotide of the invention may be engineered for expression in a replication defective retroviral vector, as discussed above. The retroviral expression construct may then be isolated and introduced into a packaging cell transduced with a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention such that the packaging cell now produces infectious viral particles containing the gene of interest. These producer cells may be administered to a subject for engineering cells in vivo and expression of the polypeptide in vivo. For overview of gene therapy, see Chapter 20, Gene Therapy and

other Molecular Genetic-based Therapeutic Approaches, (and references cited therein) in Human Molecular Genetics, T Strachan and A P Read, BIOS Scientific Publishers Ltd (1996).

Formulation and Administration

Peptides, such as the soluble form of receptor polypeptides, and agonists and antagonist peptides or small molecules, may be formulated in combination with a suitable pharmaceutical carrier. Such formulations comprise a therapeutically effective amount of the polypeptide or compound, and a pharmaceutically acceptable carrier or excipient. Such carriers include but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. Formulation should suit the mode of administration, and is well within the skill of the art. The invention further relates to pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention.

Polypeptides and other compounds of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds. Preferred forms of systemic administration of the pharmaceutical compositions include injection, typically by intravenous injection. Other injection routes, such as subcutaneous, intramuscular, or intraperitoneal, can be used. Alternative means for systemic administration include transmucosal and transdermal administration using penetrants such as bile salts or fusidic acids or other detergents. In addition, if properly formulated in enteric or encapsulated formulations, oral administration may also be possible. Administration of these compounds may also be topical and/or localized, in the form of salves, pastes, gels and the like.

The dosage range required depends on the choice of peptide, the route of administration, the nature of the formulation, the nature of the subject's condition, and the judgment of the attending practitioner. Suitable dosages, however, are in the range of 0.1-100 μg/kg of subject. Wide variations in the needed dosage, however, are to be expected in view of the variety of compounds available and the differing efficiencies of various routes of administration. For example, oral administration would be expected to require higher dosages than administration by intravenous injection. Variations in these dosage levels can be adjusted using standard empirical routines for optimization, as is well understood in the art.

Polypeptides used in treatment can also be generated endogenously in the subject, in treatment modalities often referred to as "gene therapy" as described above. Thus, for example, cells from a subject may be engineered with a polynucleotide, such as a DNA or RNA, to encode a polypeptide ex vivo, and for example, by the use of a retroviral plasmid vector. The cells are then introduced into the subject.

All publications, including but not limited to patents and patent applications, cited in this specification are herein incorporated by reference as if each individual publication were specifically and individually indicated to be incorporated by reference herein as though fully set forth.

HMACR70

```
GCAGTTCCTGAGAGAAGAACCCTGAGGAACAGACGTTCCCTCGCGGCCCTGGCACCTCCAACCCCAGATATGCTGCTGCTGCTGCTGCTG
CGTCAAGGACTCTCTTCTTGGGACTCCTTGTCTGCAAGGGAGCGCCGGGACCGTGGAGGTTGGGGTCTATACGACGACGACGACGACGAC

                                                                                    M  L  L  L  L  L  L

CCCCTGCTCTGGGGGAGGGAGAGGGTGGAAGGACAGAAGAGTAACCGGAAGGATTACTCGCTGACGATGCAGAGTTCCGTGACCGTGCAA
GGGGACGAGACCCCCTCCCTCTCCCACCTTCCTGTCTTCTCATTGGCCTTCCTAATGAGCGACTGCTACGTCTCAAGGCACTGGCACGTT

 P  L  L  W  G  R  E  R  V  E  G  Q  K  S  N  R  K  D  Y  S  L  T  M  Q  S  S  V  T  V  Q

GAGGGCATGTGTGTCCATGTGCGCTGCTCCTTCTCCTACCCAGTGGACAGCCAGACTGACTCTGACCCAGTTCATGGCTACTGGTTCCGG
CTCCCGTACACACAGGTACACGCGACGAGGAAGAGGATGGGTCACCTGTCGGTCTGACTGAGACTGGGTCAAGTACCGATGACCAAGGCC

 E  G  M  C  V  H  V  R  C  S  F  S  Y  P  V  D  S  Q  T  D  S  D  P  V  H  G  Y  W  F  R

GCAGGGAATGATATAAGCTGGAAGGCTCCAGTGGCCACAAACAACCCAGCTTGGGCAGTGCAGGAGGAAACTCGGGACCGATTCCACCTC
CGTCCCTTACTATATTCGACCTTCCGAGGTCACCGGTGTTTGTTGGGTCGAACCCGTCACGTCCTCCTTTGAGCCCTGGCTAAGGTGGAG

 A  G  N  D  I  S  W  K  A  P  V  A  T  N  N  P  A  W  A  V  Q  E  E  T  R  D  R  F  H  L

CTTGGGGACCCACAGACCAAAAATTGCACCCTGAGCATCAGAGATGCCAGAATGAGTGATGCGGGGAGATACTTCTTTCGTATGGAGAAA
GAACCCCTGGGTGTCTGGTTTTTAACGTGGGACTCGTAGTCTCTACGGTCTTACTCACTACGCCCCTCTATGAAGAAAGCATACCTCTTT

 L  G  D  P  Q  T  K  N  C  T  L  S  I  R  D  A  R  M  S  D  A  G  R  Y  F  F  R  M  E  K

GGAAATATAAAATGGAATTATAAATATGACCAGCTCTCTGTGAACGTGACATACCCTCCTCAGAACTTGACTGTGACTGTCTTCCAAGGA
CCTTTATATTTTACCTTAATATTTATACTGGTCGAGAGACACTTGCACTGTATGGGAGGAGTCTTGAACTGACACTGACAGAAGGTTCCT

 G  N  I  K  W  N  Y  K  Y  D  Q  L  S  V  N  V  T  Y  P  P  Q  N  L  T  V  T  V  F  Q  G

GAAGGCACAGCATCCACAGCTCTGGGGAACAGCTCATCTCTTTCAGTCCTAGAGGGCCAGTCTCTGCGCTTGGTCTGTGCTGTTGACAGC
CTTCCGTGTCGTAGGTGTCGAGACCCCTTGTCGAGTAGAGAAAGTCAGGATCTCCCGGTCAGAGACGCGAACCAGACACGACAACTGTCG

 E  G  T  A  S  T  A  L  G  N  S  S  S  L  S  V  L  E  G  Q  S  L  R  L  V  C  A  V  D  S

AATCCCCCTGCCAGGCTGAGCTGGACCTGGAGGAGTCTGACCCTGTACCCCTCACAGCCCTCAAACCCTCTGGTACTGGAGCTGCAAGTG
TTAGGGGGACGGTCCGACTCGACCTGGACCTCCTCAGACTGGGACATGGGGAGTGTCGGGAGTTTGGGAGACCATGACCTCGACGTTCAC

 N  P  P  A  R  L  S  W  T  W  R  S  L  T  L  Y  P  S  Q  P  S  N  P  L  V  L  E  L  Q  V

CACCTGGGGGATGAAGGGGAATTCACCTGTCGAGCTCAGAACTCTCTGGGTTCCCAGCACGTTTCCCTGAACCTCTCCCTGCAACAGGAG
GTGGACCCCCTACTTCCCCTTAAGTGGACAGCTCGAGTCTTGAGAGACCCAAGGGTCGTGCAAAGGGACTTGGAGAGGGACGTTGTCCTC

 H  L  G  D  E  G  E  F  T  C  R  A  Q  N  S  L  G  S  Q  H  V  S  L  N  L  S  L  Q  Q  E
```

```
TACACAGGCAAAATGAAGCCTGTATCAGGAGTGTTGCTGGGGGCGGTCGGGGGAACTGGAGCCACAGCCCTGGTCTTCCTCTCCTTCTGT
ATGTGTCCGTTTTACTTCGGACATAGTCCTCACAACGACCCCCGCCAGCCCCCTTGACCTCGGTGTCGGGACCAGAAGGAGAGGAAGACA

 Y  T  G  K  M  K  P  V  S  G  V  L  L  G  A  V  G  G  T  G  A  T  A  L  V  F  L  S  F  C

GTCATCTTCATTGTAGTGAGGTCCTGCAGGAAGAAATCGGCAAGACCAGCAGCGGACGTGGGAGACATAGGCATGAAGGATGCAAACACC
CAGTAGAAGTAACATCACTCCAGGACGTCCTTCTTTAGCCGTTCTGGTCGTCGCCTGCACCCTCTGTATCCGTACTTCCTACGTTTGTGG

 V  I  F  I  V  V  R  S  C  R  K  K  S  A  R  P  A  A  D  V  G  D  I  G  M  K  D  A  N  T

ATTCAGGGGCTCAGCCTCTCAGGGTAACTGGATGAGTCCTGGGCAGATGATAACCCCCGACACCATGGCCTGGCTGCCCACTCCCTCAGG
TAAGTCCCCGAGTCGGAGAGTCCCATTGACCTACTCAGGACCCGTCTACTATTGGGGGCTGTGGTACCGGACCGACGGGTGAGGGAGTCC

 I  Q  G  L  S  L  S  G   (SEQ IDNO:18)

GGAGGAAAGAGAGATCCCAGTATGCACCCCTCAGCTTTCATAAGGGGGAGCCTCAGGACCTATCCAGGTCAAGAAGCCACCAACAATGAG
CCTCCTTTCTCTCTAGGGTCATACGTGGGGAGTCGAAAGTATTCCCCCTCGGAGTCCTGGATAGGTCCAGTTCTTCGGTGGTTGTTACTC

TACTCAGAGATCAAGATCCCCAAGTAAGAAAATGCAGAGGCTCGGGCTTGTTTGAGGGTTCACGACCCCTCCAGCAAAGGAGTCTGAGGC
ATGAGTCTCTAGTTCTAGGGGTTCATTCTTTTACGTCTCCGAGCCCGAACAAACTCCCAAGTGCTGGGGAGGTCGTTTCCTCAGACTCCG

TGATTCCAGTAGAATTAGCAGCCCTCAATGCTGTGCAACAAGACATCAGAACTTATTCCTCTTGTCTAACTGAAAATGCATGCCTGATGA
ACTAAGGTCATCTTAATCGTCGGGAGTTACGACACGTTGTTCTGTAGTCTTGAATAAGGAGAACAGATTGACTTTTACGTACGGACTACT

CCAAACTCTCCCTTTCCCCATCCAATCGGTCCACACTCCCCGCCCTGGCCTCTGTACCCACCATTCTCCTCTGTACTTCTCTAAGGATGA
GGTTTGAGAGGGAAAGGGGTAGGTTAGCCAGGTGTGAGGGGCGGGACCGGAGACATGGGTGGTAAGAGGAGACATGAAGAGATTCCTACT

CTACTTTAGATTCCGAATATAGTGAGATTGTAACGTGAAAAAAAAAAAAAAAAAAAA   (SEQ IDNO:1)
GATGAAATCTAAGGCTTATATCACTCTAACATTGCACTTTTTTTTTTTTTTTTTTTT
```

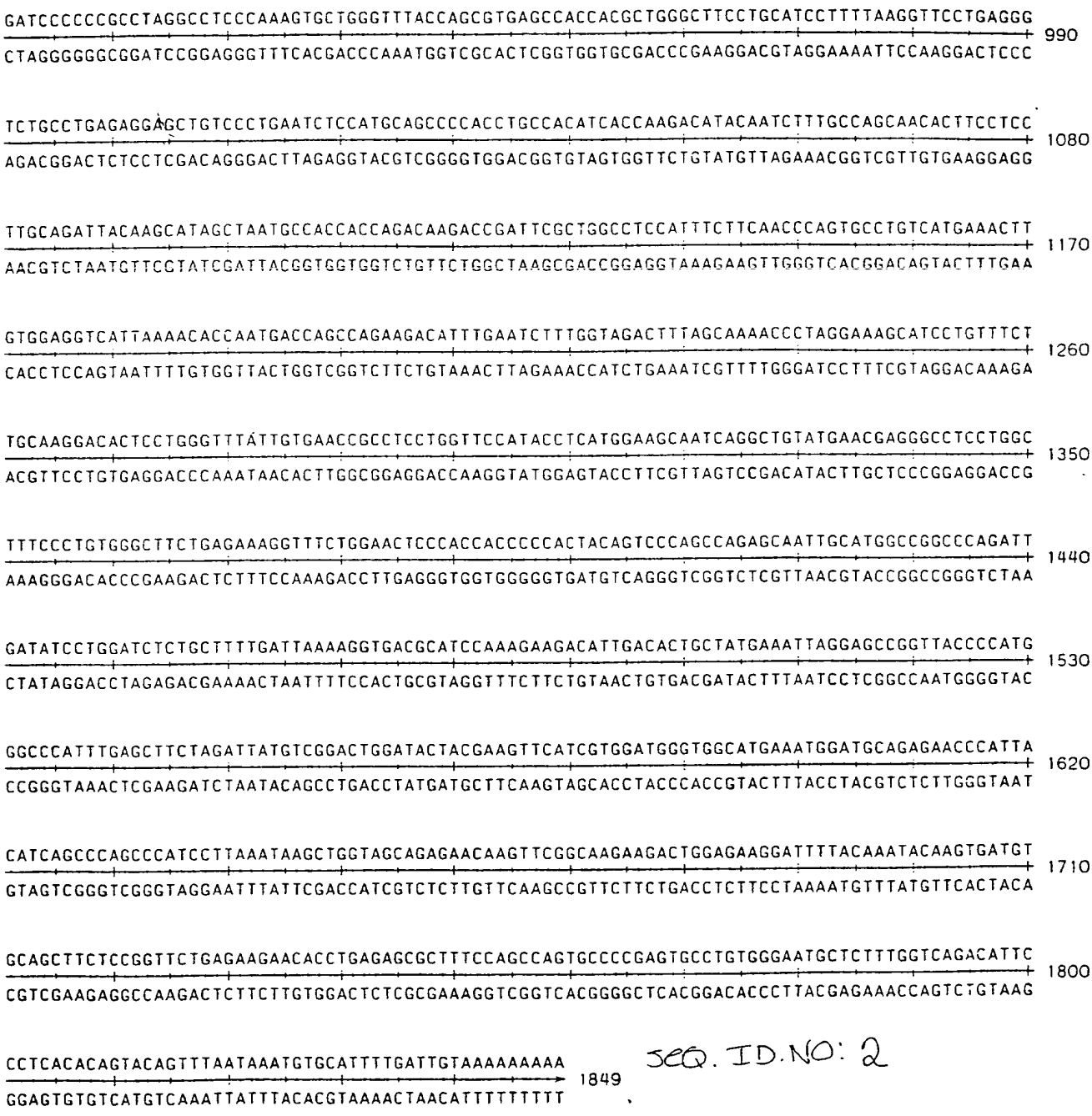
HTEDK48

```
CGGCACGAGTGGACAACCATCAGGGAGCCAGGACACAGAGGGGCAGAGCAAGTCAGCATTGGCGCCCCTTCCTCAGATCCCTATCATCTT
GCCGTGCTCACCTGTTGGTAGTCCCTCGGTCCTGTGTCTCCCCGTCTCGTTCAGTCGTAACCGCGGGGAAGGAGTCTAGGGATAGTAGAA

GGGAAACAGTAGCCCAGAGGTTCAGGAAGATGTTAACTTAAATGTTCGGGGTGCCCCAGTCTGTTCAGCATGGCTGAAATCCACACTCCG
CCCTTTGTCATCGGGTCTCCAAGTCCTTCTACAATTGAATTTACAAGCCCCACGGGGTCAGACAAGTCGTACCGACTTTAGGTGTGAGGC

                                                                    M  A  E  I  H  T  P

TATTCTTCCTTGAAGAAACTGTTATCTTTACTCAATGGCTTCGTGGCTGTGTCTGGCATCATCCTAGTTGGCCTGGGCATTGGTGGTAAA
ATAAGAAGGAACTTCTTTGACAATAGAAATGAGTTACCGAAGCACCGACACAGACCGTAGTAGGATCAACCGGACCCGTAACCACCATTT

 Y  S  S  L  K  K  L  L  S  L  L  N  G  F  V  A  V  S  G  I  I  L  V  G  L  G  I  G  G  K

TGTGGAGGGGCCTCTCTGACGAATGTCCTCGGGCTGTCCTCCGCATACCTCCTTCACGTTGGCAACCTGTGCCTGGTGATGGGATGCATC
ACACCTCCCCGGAGAGACTGCTTACAGGAGCCCGACAGGAGGCGTATGGAGGAAGTGCAACCGTTGGACACGGACCACTACCCTACGTAG

 C  G  G  A  S  L  T  N  V  L  G  L  S  S  A  Y  L  L  H  V  G  N  L  C  L  V  M  G  C  I

ACGGTACTGCTTGGCTGTGCCGGGTGGTATGGAGCGACTAAAGAGAGCAGAGGCACGCTCTTGTTTTGCATCCTGTCAATGGTTATTGTC
TGCCATGACGAACCGACACGGCCCACCATACCTCGCTGATTTCTCTCGTCTCCGTGCGAGAACAAAACGTAGGACAGTTACCAATAACAG

 T  V  L  L  G  C  A  G  W  Y  G  A  T  K  E  S  R  G  T  L  L  F  C  I  L  S  M  V  I  V

CTCATCATGGAAGTTACAGCTGCCACAGTGGTCCTTCTTTTCTTTCCAATTGTTGGAGATGTGGCCTTGGAACACACCTTCGTGACCCTG
GAGTAGTACCTTCAATGTCGACGGTGTCACCAGGAAGAAAAGAAAGGTTAACAACCTCTACACCGGAACCTTGTGTGGAAGCACTGGGAC

 L  I  M  E  V  T  A  A  T  V  V  L  L  F  F  P  I  V  G  D  V  A  L  E  H  T  F  V  T  L

AGGAAGAATTACAGAGGTTACAACGAGCCAGACGACTATTCTACACAGTGGAACTTGGTCATGGAGAAGCTAAAGTGCTGTGGGGTGAAT
TCCTTCTTAATGTCTCCAATGTTGCTCGGTCTGCTGATAAGATGTGTCACCTTGAACCAGTACCTCTTCGATTTCACGACACCCCACTTA

 R  K  N  Y  R  G  Y  N  E  P  D  D  Y  S  T  Q  W  N  L  V  M  E  K  L  K  C  C  G  V  N

AACTACACAGATTTTTCTGGCTCTTCCTTCGAAATGACAACGGGCCACACCTACCCCAGGAGTTGCTGTAAATCCATCGGAAGTGTGTCC
TTGATGTGTCTAAAAAGACCGAGAAGGAAGCTTTACTGTTGCCCGGTGTGGATGGGGTCCTCAACGACATTTAGGTAGCCTTCACACAGG

 N  Y  T  D  F  S  G  S  S  F  E  M  T  T  G  H  T  Y  P  R  S  C  C  K  S  I  G  S  V  S

TGTGACGGACGCGATGTGTCTCCAAACGTCATCCACCAGAAGGGCTGTTTCCATAAACTCCTAAAAATCACCAAGACTCAGAGCTTCACC
ACACTGCCTGCGCTACACAGAGGTTTGCAGTAGGTGGTCTTCCCGACAAAGGTATTTGAGGATTTTTAGTGGTTCTGAGTCTCGAAGTGG

 C  D  G  R  D  V  S  P  N  V  I  H  Q  K  G  C  F  H  K  L  L  K  I  T  K  T  Q  S  F  T

CTGAGTGGGAGCTCTCTGGGAGCTGCAGTGATACAGAGGTGGGGGTCTCGCTATGTTGCGCAGGCTGGTCTTGAACTGCTGGCCTAAAGC
GACTCACCCTCGAGAGACCCTCGACGTCACTATGTCTCCACCCCCAGAGCGATACAACGCGTCCGACCAGAACTTGACGACCGGATTTCG

 L  S  G  S  S  L  G  A  A  V  I  Q  R  W  G  S  R  Y  V  A  Q  A  G  L  E  L  L  A
```

(SEQ ID NO:3)

(SEQ ID NO: 19)

HTEDK48



SEQ. ID. NO: 2

HPWAE25

```
GGCGTCCCTCTGCCTGCCCACTCAGTGGCAACACCCGGGAGCTGTTTTGTCCTTTGTGGAGCCTCAGCAGTTCCCTCTTTCAGAACTCAC
CCGCAGGGAGACGGACGGGTGAGTCACCGTTGTGGGCCCTCGACAAAACAGGAAACACCTCGGAGTCGTCAAGGGAGAAAGTCTTGAGTG                 90

TGCCAAGAGCCCTGAACAGGAGCCACCATGCAGTGCTTCAGCTTCATTAAGACCATGATGATCCTCTTCAATTTGCTCATCTTTCTGTGT
ACGGTTCTCGGGACTTGTCCTCGGTGGTACGTCACGAAGTCGAAGTAATTCTGGTACTACTAGGAGAAGTTAAACGAGTAGAAAGACACA                180

                                M  Q  C  F  S  F  I  K  T  M  M  I  L  F  N  L  L  I  F  L  C

GGTGCAGCCCTGTTGGCAGTGGGCATCTGGGTGTCAATCGATGGGGCATCCTTTCTGAAGATCTTCGGGCCACTGTCGTCCAGTGCCATG
CCACGTCGGGACAACCGTCACCCGTAGACCCACAGTTAGCTACCCCGTAGGAAAGACTTCTAGAAGCCCGGTGACAGCAGGTCACGGTAC                270

 G  A  A  L  L  A  V  G  I  W  V  S  I  D  G  A  S  F  L  K  I  F  G  P  L  S  S  S  A  M

CAGTTTGTCAACGTGGGCTACTTCCTCATCGCAGCCGGCGTTGTGGTCTTTGCTCTTGGTTTCCTGGGCTGCTATGGTGCTAAGACTGAG
GTCAAACAGTTGCACCCGATGAAGGAGTAGCGTCGGCCGCAACACCAGAAACGAGAACCAAAGGACCCGACGATACCACGATTCTGACTC                360

 Q  F  V  N  V  G  Y  F  L  I  A  A  G  V  V  V  F  A  L  G  F  L  G  C  Y  G  A  K  T  E

AGCAAGTGTGCCCTCGTGACGTTCTTCTTCATCCTCCTCCTCATCTTCATTGCTGAGGTTGCAGCTGCTGTGGTCGCCTTGGTGTACACC
TCGTTCACACGGGAGCACTGCAAGAAGAAGTAGGAGGAGGAGTAGAAGTAACGACTCCAACGTCGACGACACCAGCGGAACCACATGTGG                450

 S  K  C  A  L  V  T  F  F  F  I  L  L  L  I  F  I  A  E  V  A  A  A  V  V  A  L  V  Y  T

ACAATGGCTGAGCACTTCCTGACGTTGCTGGTAGTGCCTGCCATCAAGAAAGATTATGGTTCCCAGGAAGACTTCACTCAAGTGTGGAAC
TGTTACCGACTCGTGAAGGACTGCAACGACCATCACGGACGGTAGTTCTTTCTAATACCAAGGGTCCTTCTGAAGTGAGTTCACACCTTG                540

 T  M  A  E  H  F  L  T  L  L  V  V  P  A  I  K  K  D  Y  G  S  Q  E  D  F  T  Q  V  W  N

ACCACCATGAAAGGGCTCAAGTGCTGTGGCTTCACCAACTATACGGATTTTGAGGACTCACCCTACTTCAAAGAGAACAGTGCCTTTCCC
TGGTGGTACTTTCCCGAGTTCACGACACCGAAGTGGTTGATATGCCTAAAACTCCTGAGTGGGATGAAGTTTCTCTTGTCACGGAAAGGG                630

 T  T  M  K  G  L  K  C  C  G  F  T  N  Y  T  D  F  E  D  S  P  Y  F  K  E  N  S  A  F  P

CCATTCTGTTGCAATGACAACGTCACCAACACAGCCAATGAAACCTGCACCAAGCAAAAGGCTCACGACCAAAAAGTAGAGGGTTGCTTC
GGTAAGACAACGTTACTGTTGCAGTGGTTGTGTCGGTTACTTTGGACGTGGTTCGTTTTCCGAGTGCTGGTTTTTCATCTCCCAACGAAG                720

 P  F  C  C  N  D  N  V  T  N  T  A  N  E  T  C  T  K  Q  K  A  H  D  Q  K  V  E  G  C  F

AATCAGCTTTTGTATGACATCCGAACTAATGCAGTCACCGTGGGTGGTGTGGCAGCTGGAATTGGGGGCCTCGAGCTGGCTGCCATGAAT
TTAGTCGAAAACATACTGTAGGCTTGATTACGTCAGTGGCACCCACCACACCGTCGACCTTAACCCCCGGAGCTCGACCGACGGTACTTA                810

 N  Q  L  L  Y  D  I  R  T  N  A  V  T  V  G  G  V  A  A  G  I  G  G  L  E  L  A  A  M  N

TGTGTCCATGTATCTGTACTGCAATCTACAATAAGTCCACTTCTGCCTCTGCCACTACTGCTGCCACATGGGAACTGTGAAGAGGCACCC
ACACAGGTACATAGACATGACGTTAGATGTTATTCAGGTGAAGACGGAGACGGTGATGACGACGGTGTACCCTTGACACTTCTCCGTGGG                900

 C  V  H  V  S  V  L  Q  S  T  I  S  P  L  L  P  L  P  L  L  L  P  H  G  N  C  E  E  A  P
```

```
TGGCAAGCAGCAGTGATTGGGGGAGGGGACAGGATCTAACAATGTCACTTGGGCCAGAATGGACCTGCCCTTTCTGCTCCAGACTTGGGG
ACCGTTCGTCGTCACTAACCCCCTCCCCTGTCCTAGATTGTTACAGTGAACCCGGTCTTACCTGGACGGGAAAGACGAGGTCTGAACCCC
```

990

W Q A A V I G G G D R I (SEQ ID NO: 20)

```
CTAGATAGGGACCACTCCTTTTAGGCGATGCCTGACTTTCCTTCCATTGGTGGGTGGATGGGTGGGGGGCATTCCAGAGCCTCTAAGGTA
GATCTATCCCTGGTGAGGAAAATCCGCTACGGACTGAAAGGAAGGTAACCACCCACCTACCCACCCCCCGTAAGGTCTCGGAGATTCCAT
```

1080

```
GCCAGTTCTGTTGCCCATTCCCCCAGTCTATTAAACCCTTGATATGCCCCCTAGGCCTAGTGGTGATCCCAGTGCTCTACTGGGGGATGA
CGGTCAAGACAACGGGTAAGGGGGTCAGATAATTTGGGAACTATACGGGGGATCCGGATCACCACTAGGGTCACGAGATGACCCCCTACT
```

1170

```
GAGAAAGGCATTTTATAGCCTGGGCATAAGTGAAATCAGCAGAGCCTCTGGGTGGATGTGTAGAAGGCACTTCAAAATGCATAAACCTGT
CTCTTTCCGTAAAATATCGGACCCGTATTCACTTTAGTCGTCTCGGAGACCCACCTACACATCTTCCGTGAAGTTTTACGTATTTGGACA
```

1260

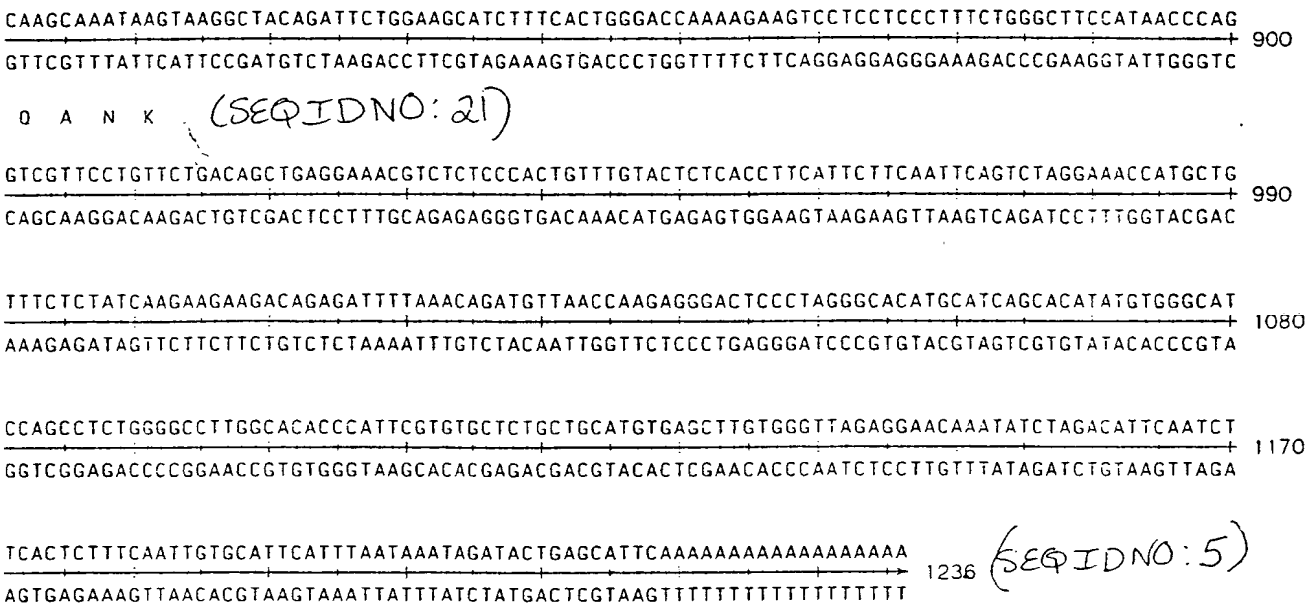
```
TACAATGTTAAAAAAAAAAAAAAAAAA
ATGTTACAATTTTTTTTTTTTTTTTTT
```

1288 (SEQ ID NO: 4)

HTPEF86

```
AAAAAAAACAAGGTCCCCACAGCAAAGAAAAGGAATAGGATCAAGAGATACGTGGCTGCTGGCAGAGCAAGCATGAATTCGATGACTTCA
---------+---------+---------+---------+---------+---------+---------+---------+---------+ 90
TTTTTTTTGTTCCAGGGGTGTCGTTTCTTTTCCTTATCCTAGTTCTCTATGCACCGACGACCGTCTCGTTCGTACTTAAGCTACTGAAGT

                                                                          M  N  S  M  T  S

GCAGTTCCGGTGGCCAATTCTGTGTTGGTGGTGGCACCCCACAATGGTTATCCTGTGACCCCAGGAATTATGTCTCACGTGCCCCTGTAT
---------+---------+---------+---------+---------+---------+---------+---------+---------+ 180
CGTCAAGGCCACCGGTTAAGACACAACCACCACCGTGGGGTGTTACCAATAGGACACTGGGGTCCTTAATACAGAGTGCACGGGGACATA

 A  V  P  V  A  N  S  V  L  V  V  A  P  H  N  G  Y  P  V  T  P  G  I  M  S  H  V  P  L  Y

CCAAACAGCCAGCCGCAAGTCCACCTAGTTCCTGGGAACCCACCTAGTTTGGTGTCGAATGTGAATGGGCAGCCTGTGCAGAAAGCTCTG
---------+---------+---------+---------+---------+---------+---------+---------+---------+ 270
GGTTTGTCGGTCGGCGTTCAGGTGGATCAAGGACCCTTGGGTGGATCAAACCACAGCTTACACTTACCCGTCGGACACGTCTTTCGAGAC

 P  N  S  Q  P  Q  V  H  L  V  P  G  N  P  P  S  L  V  S  N  V  N  G  Q  P  V  Q  K  A  L

AAAGAAGGCAAAACCTTGGGGGCCATCCAGATCATCATTGGCCTGGCTCACATCGGCCTCGGCTCCATCATGGCGACGGTTCTCGTAGGG
---------+---------+---------+---------+---------+---------+---------+---------+---------+ 360
TTTCTTCCGTTTTGGAACCCCCGGTAGGTCTAGTAGTAACCGGACCGAGTGTAGCCGGAGCCGAGGTAGTACCGCTGCCAAGAGCATCCC

 K  E  G  K  T  L  G  A  I  Q  I  I  I  G  L  A  H  I  G  L  G  S  I  M  A  T  V  L  V  G

GAATACCTGTCTATTTCATTCTACGGAGGCTTTCCCTTCTGGGGAGGCTTGTGGTTTATCATTTCAGGATCTCTCTCCGTGGCAGCAGAA
---------+---------+---------+---------+---------+---------+---------+---------+---------+ 450
CTTATGGACAGATAAAGTAAGATGCCTCCGAAAGGGAAGACCCCTCCGAACACCAAATAGTAAAGTCCTAGAGAGAGGCACCGTCGTCTT

 E  Y  L  S  I  S  F  Y  G  G  F  P  F  W  G  G  L  W  F  I  I  S  G  S  L  S  V  A  A  E

AATCAGCCATATTCTTATTGCCTGCTGTCTGGCAGTTTGGGCTTGAACATCGTCAGTGCAATCTGCTCTGCAGTTGGAGTCATACTCTTC
---------+---------+---------+---------+---------+---------+---------+---------+---------+ 540
TTAGTCGGTATAAGAATAACGGACGACAGACCGTCAAACCCGAACTTGTAGCAGTCACGTTAGACGAGACGTCAACCTCAGTATGAGAAG

 N  Q  P  Y  S  Y  C  L  L  S  G  S  L  G  L  N  I  V  S  A  I  C  S  A  V  G  V  I  L  F

ATCACAGATCTAAGTATTCCCCACCCATATGCCTACCCCGACTATTATCCTTACGCCTGGGGTGTGAACCCTGGAATGGCGATTTCTGGC
---------+---------+---------+---------+---------+---------+---------+---------+---------+ 630
TAGTGTCTAGATTCATAAGGGGTGGGTATACGGATGGGGCTGATAATAGGAATGCGGACCCCACACTTGGGACCTTACCGCTAAAGACCG

 I  T  D  L  S  I  P  H  P  Y  A  Y  P  D  Y  Y  P  Y  A  W  G  V  N  P  G  M  A  I  S  G

GTGCTGCTGGTCTTCTGCCTCCTGGAGTTTGGCATCGCATGCGCATCTTCCCACTTTGGCTGCCAGTTGGTCTGCTGTCAATCAAGCAAT
---------+---------+---------+---------+---------+---------+---------+---------+---------+ 720
CACGACGACCAGAAGACGGAGGACCTCAAACCGTAGCGTACGCGTAGAAGGGTGAAACCGACGGTCAACCAGACGACAGTTAGTTCGTTA

 V  L  L  V  F  C  L  L  E  F  G  I  A  C  A  S  S  H  F  G  C  Q  L  V  C  C  Q  S  S  N

GTGAGTGTCATCTATCCAAACATCTATGCAGCAAACCCAGTGATCACCCCAGAACCGGTGACCTCACCACCAAGTTATTCCAGTGAGATC
---------+---------+---------+---------+---------+---------+---------+---------+---------+ 810
CACTCACAGTAGATAGGTTTGTAGATACGTCGTTTGGGTCACTAGTGGGGTCTTGGCCACTGGAGTGGTGGTTCAATAAGGTCACTCTAG

 V  S  V  I  Y  P  N  I  Y  A  A  N  P  V  I  T  P  E  P  V  T  S  P  P  S  Y  S  S  E  I
```

HTPEF86


CAAGCAAATAAGTAAGGCTACAGATTCTGGAAGCATCTTTCACTGGGACCAAAAGAAGTCCTCCTCCCTTTCTGGGCTTCCATAACCCAG
900
GTTCGTTTATTCATTCCGATGTCTAAGACCTTCGTAGAAAGTGACCCTGGTTTTCTTCAGGAGGAGGGAAAGACCCGAAGGTATTGGGTC
Q A N K (SEQIDNO:21)
GTCGTTCCTGTTCTGACAGCTGAGGAAACGTCTCTCCCACTGTTTGTACTCTCACCTTCATTCTTCAATTCAGTCTAGGAAACCATGCTG
990
CAGCAAGGACAAGACTGTCGACTCCTTTGCAGAGAGGGTGACAAACATGAGAGTGGAAGTAAGAAGTTAAGTCAGATCCTTTGGTACGAC
TTTCTCTATCAAGAAGAAGACAGAGATTTTAAACAGATGTTAACCAAGAGGGACTCCCTAGGGCACATGCATCAGCACATATGTGGGCAT
1080
AAAGAGATAGTTCTTCTTCTGTCTCTAAAATTTGTCTACAATTGGTTCTCCCTGAGGGATCCCGTGTACGTAGTCGTGTATACACCCGTA
CCAGCCTCTGGGGCCTTGGCACACCCATTCGTGTGCTCTGCTGCATGTGAGCTTGTGGGTTAGAGGAACAAATATCTAGACATTCAATCT
1170
GGTCGGAGACCCCGGAACCGTGTGGGTAAGCACACGAGACGACGTACACTCGAACACCCAATCTCCTTGTTTATAGATCTGTAAGTTAGA
TCACTCTTTCAATTGTGCATTCATTTAATAAATAGATACTGAGCATTCAAAAAAAAAAAAAAAAAA
1236 (SEQIDNO:5)
AGTGAGAAAGTTAACACGTAAGTAAATTATTTATCTATGACTCGTAAGTTTTTTTTTTTTTTTTTT

HSBBF02

```
CACGAGCAGGGTCTCGGGCTAGTCATGGCGTCCCCGTCTCGGAGACTGCAGACTAAACCAGTCATTACTTGTTTCAAGAGCGTTCTGCTA
                                                                                            90
GTGCTCGTCCCAGAGCCCGATCAGTACCGCAGGGGCAGAGCCTCTGACGTCTGATTTGGTCAGTAATGAACAAAGTTCTCGCAAGACGAT

                    M  A  S  P  S  R  R  L  Q  T  K  P  V  I  T  C  F  K  S  V  L  L

ATCTACACTTTTATTTTCTGGATCACTGGCGTTATCCTTCTTGCAGTTGGCATTTGGGGCAAGGTGAGCCTGGAGAATTACTTTTCTCTT
                                                                                           180
TAGATGTGAAAATAAAAGACCTAGTGACCGCAATAGGAAGAACGTCAACCGTAAACCCCGTTCCACTCGGACCTCTTAATGAAAAGAGAA

 I  Y  T  F  I  F  W  I  T  G  V  I  L  L  A  V  G  I  W  G  K  V  S  L  E  N  Y  F  S  L

TTAAATGAGAAGGCCACCAATGTCCCCTTCGTGCTCATTGCTACTGGTACCGTCATTATTCTTTTGGGCACCTTTGGTTGTTTTGCTACC
                                                                                           270
AATTTACTCTTCCGGTGGTTACAGGGGAAGCACGAGTAACGATGACCATGGCAGTAATAAGAAAACCCGTGGAAACCAACAAAACGATGG

 L  N  E  K  A  T  N  V  P  F  V  L  I  A  T  G  T  V  I  I  L  L  G  T  F  G  C  F  A  T

TGCCGAGCTTCTGCATGGATGCTAAAACTGTATGCAATGTTTCTGACTCTCGTTTTTTTGGTCGAACTGGTCGCTGCCATCGTAGGATTT
                                                                                           360
ACGGCTCGAAGACGTACCTACGATTTTGACATACGTTACAAAGACTGAGAGCAAAAAAACCAGCTTGACCAGCGACGGTAGCATCCTAAA

 C  R  A  S  A  W  M  L  K  L  Y  A  M  F  L  T  L  V  F  L  V  E  L  V  A  A  I  V  G  F

GTTTTCAGACATGAGATTAAGAACAGCTTTAAGAATAATTATGAGAAGGCTTTGAAGCAGTATAACTCTACAGGAGATTATAGAAGCCAT
                                                                                           450
CAAAAGTCTGTACTCTAATTCTTGTCGAAATTCTTATTAATACTCTTCCGAAACTTCGTCATATTGAGATGTCCTCTAATATCTTCGGTA

 V  F  R  H  E  I  K  N  S  F  K  N  N  Y  E  K  A  L  K  Q  Y  N  S  T  G  D  Y  R  S  H

GCAGTAGACAAGATCCAAAATACGTTGCATTGTTGTGGTGTCACCGATTATAGAGATTGGACAGATACTAATTATTACTCAGAAAAAGGA
                                                                                           540
CGTCATCTGTTCTAGGTTTTATGCAACGTAACAACACCACAGTGGCTAATATCTCTAACCTGTCTATGATTAATAATGAGTCTTTTTCCT

 A  V  D  K  I  Q  N  T  L  H  C  C  G  V  T  D  Y  R  D  W  T  D  T  N  Y  Y  S  E  K  G

TTTCCTAAGAGTTGCTGTAAACTTGAAGATTGTACTCCACAGAGAGATGCAGACAAAGTAAACAATGAAGGTTGTTTTATAAAGGTGATG
                                                                                           630
AAAGGATTCTCAACGACATTTGAACTTCTAACATGAGGTGTCTCTCTACGTCTGTTTCATTTGTTACTTCCAACAAAATATTTCCACTAC

 F  P  K  S  C  C  K  L  E  D  C  T  P  Q  R  D  A  D  K  V  N  N  E  G  C  F  I  K  V  M

ACCATTATAGAGTCAGAAATGGGAGTCGTTGCAGGAATTTCCTTTGGAGTTGCTTGCTTCCAACTGATTGGAATCTTTCTCGCCTACTGC
                                                                                           720
TGGTAATATCTCAGTCTTTACCCTCAGCAACGTCCTTAAAGGAAACCTCAACGAACGAAGGTTGACTAACCTTAGAAAGAGCGGATGACG

 T  I  I  E  S  E  M  G  V  V  A  G  I  S  F  G  V  A  C  F  Q  L  I  G  I  F  L  A  Y  C

CTCTCTCGTGCCATAACAAATAACCAGTATGAGATAGTGTAACCCAATGTATCTGTGGGCCTATTCCTCTCTACCTTTAAGGACATTTAG
                                                                                           810
GAGAGAGCACGGTATTGTTTATTGGTCATACTCTATCACATTGGGTTACATAGACACCCGGATAAGGAGAGATGGAAATTCCTGTAAATC

 L  S  R  A  I  T  N  N  Q  Y  E  I  V .  (SEQ ID NO: 22)

GGTCCCCCCTGTGAATTAGAAAGTTGCTTGGCTGGAGAACTGACAACACTACTTACTGATAGACCAAAAAACTACACCAGTAGGTTGATT
                                                                                           900
CCAGGGGGGACACTTAATCTTTCAACGAACCGACCTCTTGACTGTTGTGATGAATGACTATCTGGTTTTTTGATGTGGTCATCCAACTAA
```

HSBBF02

CAATCAAGATGTATGTAGACCTAAAACTACACCAATAGGCTGATTCAATCAAGATCCGTGCTCGCAGTGGGCTGATTCAATCAAGATGTA 990
GTTAGTTCTACATACATCTGGATTTTGATGTGGTTATCCGACTAAGTTAGTTCTAGGCACGAGCGTCACCCGACTAAGTTAGTTCTACAT

TGTTTGCTATGTTCTAAGTCCACCTTCTATCCCATTCATGTTAGATCGTTGAAACCCTGTATCCCTCTGAAACACTGGAAGAGCTAGTAA 1080
ACAAACGATACAAGATTCAGGTGGAAGATAGGGTAAGTACAATCTAGCAACTTTGGGACATAGGGAGACTTTGTGACCTTCTCGATCATT

ATTGTAAATGAAGTAAAAAAAAAAAAAAAAAAAAA 1115 (SEQ ID NO:6)
TAACATTTACTTCATTTTTTTTTTTTTTTTTTTTT

HLTAH80

CACGAGCATTGCCGCTCTCTCGGTGAGCGCAGCCCCGCTCTCCGGGCCGGGCCTTCGCGGGCCACCGGCGCCATGGGCCAGTGCGGCATC 90
GTGCTCGTAACGGCGAGAGAGCCACTCGCGTCGGGGCGAGAGGCCCGGCCCGGAAGCGCCCGGTGGCCGCGGTACCCGGTCACGCCGTAG

M G Q C G I

ACCTCCTCCAAGACCGTGCTGGTCTTTCTCAACCTCATCTTCTGGGGGGCAGCTGGCATTTTATGCTATGTGGGAGCCTATGTCTTCATC 180
TGGAGGAGGTTCTGGCACGACCAGAAAGAGTTGGAGTAGAAGACCCCCCGTCGACCGTAAAATACGATACACCCTCGGATACAGAAGTAG

T S S K T V L V F L N L I F W G A A G I L C Y V G A Y V F I

ACTTATGATGACTATGACCACTTCTTTGAAGATGTGTACACGCTCATCCCTGCTGTAGTGATCATAGCTGTAGGAGCCCTGCTTTTCATC 270
TGAATACTACTGATACTGGTGAAGAAACTTCTACACATGTGCGAGTAGGGACGACATCACTAGTATCGACATCCTCGGGACGAAAAGTAG

T Y D D Y D H F F E D V Y T L I P A V V I I A V G A L L F I

ATTGGGCTAATTGGCTGCTGTGCCACAATCCGGGAAAGTCGCTGTGGACTTGCCACGTTTGTCATCATCCTGCTCTTGGTTTTTGTCACA 360
TAACCCGATTAACCGACGACACGGTGTTAGGCCCTTTCAGCGACACCTGAACGGTGCAAACAGTAGTAGGACGAGAACCAAAAACAGTGT

I G L I G C C A T I R E S R C G L A T F V I I L L L V F V T

GAAGTTGTTGTAGTGGTTTTGGGATATGTTTACAGAGCAAAGGTGGAAAATGAGGTTGATCGCAGCATTCAGAAAGTGTATAAGACCTAC 450
CTTCAACAACATCACCAAAACCCTATACAAATGTCTCGTTTCCACCTTTTACTCCAACTAGCGTCGTAAGTCTTTCACATATTCTGGATG

E V V V V V L G Y V Y R A K V E N E V D R S I Q K V Y K T Y

AATGGAACCAACCCTGATGCTGCTAGCCGGGCTATTGATTATGTACAGAGACAGCTGCATTGTTGTGGAATTCACAACTACTCAGACTGG 540
TTACCTTGGTTGGGACTACGACGATCGGCCCGATAACTAATACATGTCTCTGTCGACGTAACAACACCTTAAGTGTTGATGAGTCTGACC

N G T N P D A A S R A I D Y V Q R Q L H C C G I H N Y S D W

GAAAATACAGATTGGTTCAAAGAAACCAAAAACCAGAGTGTCCCTCTTAGCTGCTGCAGAGAGACTGCCAGCAATTGTAATGGCAGCCTG 630
CTTTTATGTCTAACCAAGTTTCTTTGGTTTTTGGTCTCACAGGGAGAATCGACGACGTCTCTCTGACGGTCGTTAACATTACCGTCGGAC

E N T D W F K E T K N Q S V P L S C C R E T A S N C N G S L

GCCCACCCTTCCGACCTCTATGCTGAGGGGTGTGAGGCTCTAGTTGTGAAGAAGCTACAAGAAATCATGATGCATGTGATCTGGGCCGCA 720
CGGGTGGGAAGGCTGGAGATACGACTCCCCACACTCCGAGATCAACACTTCTTCGATGTTCTTTAGTACTACGTACACTAGACCCGGCGT

A H P S D L Y A E G C E A L V V K K L Q E I M M H V I W A A

CTGGCATTTGCAGCTATTCAGCTGCTGGGCATGCTGTGTGCTTGCATCGTGTTGTGCAGAAGGAGTAGAGATCCTGCTTACGAGCTCTTC 810
GACCGTAAACGTCGATAAGTCGACGACCCGTACGACACACGAACGTAGCACAACACGTCTTCCTCATCTCTAGGACGAATGCTCGAGAAG

L A F A A I Q L L G M L C A C I V L C R R S R D P A Y E L F

HLTAH8O

```
ATCACTGGCGGAACCTATGCATAGTTGACAATCTCAAGCCTGAGCTTTTTGGTCTTGTTCTGATTTGGAAGGTGAATTGAGCAGGTCTGC
TAGTGACCGCCTTGGATACGTATCAACTGTTAGAGTTCGGACTCGAAAAACCAGAACAAGACTAAACCTTCCACTTAACTCGTCCAGACG
```

900

I T G G T Y A . (SEQ ID NO: 23)

```
TGCTGTTGGCCTCTGGAGTTCATTTAGTTAAAGCACATGTACACTGGTGTTGGACAGAGCAGCTTGGCTTTTCATGTGCCCACCTACTTA
ACGACAACCGGAGACCTCAAGTAAATCAATTTCGTGTACATGTGACCACAACCTGTCTCGTCGAACCGAAAAGTACACGGGTGGATGAAT
```

990

```
CCTACTACCTGCGACTTTCTTTTTCCTTGTTCTAGCTGACTCTTCATGCCCCTAAGATTTTAAGTACGATGGTGAACGTTCTAATTTCAG
GGATGATGGACGCTGAAAGAAAAAGGAACAAGATCGACTGAGAAGTACGGGGATTCTAAAATTCATGCTACCACTTGCAAGATTAAAGTC
```

1080

```
AACCAATTGCGAGTCATGTAGTGTGGTAGAATTAAAGGAGGACACGAGCCTGCTTCTGTTACCTCCAAGTGGTAACAGGACTGATGCCGA
TTGGTTAACGCTCAGTACATCACACCATCTTAATTTCCTCCTGTGCTCGGACGAAGACAATGGAGGTTCACCATTGTCCTGACTACGGCT
```

1170

```
AATGTCACCAGGTCCTTTCAGTCTTCACAGTGGAGAACTCTTGGCCAAAGGTTTTTGGGGGGAGGAGGAGGAAACCAGCTTTCTGGTTAA
TTACAGTGGTCCAGGAAAGTCAGAAGTGTCACCTCTTGAGAACCGGTTTCCAAAAACCCCCCTCCTCCTCCTTTGGTCGAAAGACCAATT
```

1260

```
GGTTAACACCAGATGGTGCCCCTCATTGGTGTCCTTTTAAAAAATATTTACTGTAGTCCAATAAGATAGCAGCTGTACAAAATGACTAAA
CCAATTGTGGTCTACCACGGGGAGTAACCACAGGAAAATTTTTTATAAATGACATCAGGTTATTCTATCGTCGACATGTTTTACTGATTT
```

1350

```
ATAGATTGTAGGATCATATGGCGTATATCTTGGTTCATCTTCAAAATCAGAGACTGAGCTTTGAAACTAGTGGTTTTTAATCAAAGTTGG
TATCTAACATCCTAGTATACCGCATATAGAACCAAGTAGAAGTTTTAGTCTCTGACTCGAAACTTTGATCACCAAAAATTAGTTTCAACC
```

1440

```
CTTTATAGGAGGAGTATAATGTATGCACTACTGTTTTAAAAGAATTAGTGTGAGTGTGTTTTTGTATGAATGAGCCCATTCATGGTAAGT
GAAATATCCTCCTCATATTACATACGTGATGACAAAATTTTCTTAATCACACTCACACAAAAACATACTTACTCGGGTAAGTACCATTCA
```

1530

```
CTTAAGCTTGTTGGAAATAATGTACCCATGTAGACTAGCAAAATAGTATGTAGATGTGATCTCAGTTGTAAATAGAAAAATCTAATTCAA
GAATTCGAACAACCTTTATTACATGGGTACATCTGATCGTTTTATCATACATCTACACTAGAGTCAACATTTATCTTTTTAGATTAAGTT
```

1620

```
TAAACTCTGTATCAGCCCCCAACAAAAAAAAAAAAAAAAAAA
ATTTGAGACATAGTCGGGGGTTGTTTTTTTTTTTTTTTTTTT
```

1662 (SEQ ID NO: 7)

HTPBA27

```
CACGAGCGCAGAGCTTGGGGCTTCCTTGGTCGCACCCACCACCTGCCTGCCCACTGGTCAGCCTTCAGGGACCCTGAGCACCGCCTGGTC
                                                                                          90
GTGCTCGCGTCTCGAACCCCGAAGGAACCAGCGTGGGTGGTGGACGGACGGGTGACCAGTCGGAAGTCCCTGGGACTCGTGGCGGACCAG

TCTTTCCTGTGGCCAGCCCAGAACTGAAGCGCTGCGGCATGGCGCGCGCCTGCCTCCAGGCCGTCAAGTACCTCATGTTCGCCTTCAACC
                                                                                         180
AGAAAGGACACCGGTCGGGTCTTGACTTCGCGACGCCGTACCGCGCGCGGACGGAGGTCCGGCAGTTCATGGAGTACAAGCGGAAGTTGG

                                        M  A  R  A  C  L  Q  A  V  K  Y  L  M  F  A  F  N

TGTTCTTCTGGCTGGGAGGCTGTGGCGTGCTGGGTGTCGGCATCTGGCTGGCCGCCACACAGGGGAGCTTCGCCACGCTGTCTTCTTCCT
                                                                                         270
ACAAGAAGACCGACCCTCCGACACCGCACGACCCACAGCCGTAGACCGACCGGCGGTGTGTCCCCTCGAAGCGGTGCGACAGAAGAAGGA

L  F  F  W  L  G  G  C  G  V  L  G  V  G  I  W  L  A  A  T  Q  G  S  F  A  T  L  S  S  S

TCCCGTCCCTGTCGGCTGCCAACCTGCTCATCATCACCGGCGCCTTTGTCATGGCCATCGGCTTCGTGGGCTGCCTGGGTGCCATCAAGG
                                                                                         360
AGGGCAGGGACAGCCGACGGTTGGACGAGTAGTAGTGGCCGCGGAAACAGTACCGGTAGCCGAAGCACCCGACGGACCCACGGTAGTTCC

F  P  S  L  S  A  A  N  L  L  I  I  T  G  A  F  V  M  A  I  G  F  V  G  C  L  G  A  I  K

AGAACAAGTGCCTCCTGCTCACTTTCTTCCTGCTGCTGCTGCTGGTGTTCCTGCTGGAGGCCACCATCGCCATCCTCTTCTTCGCCTACA
                                                                                         450
TCTTGTTCACGGAGGACGAGTGAAAGAAGGACGACGACGACGACCACAAGGACGACCTCCGGTGGTAGCGGTAGGAGAAGAAGCGGATGT

E  N  K  C  L  L  L  T  F  F  L  L  L  L  L  V  F  L  L  E  A  T  I  A  I  L  F  F  A  Y

CGGACAAGATTGACAGGTATGCCCAGCAAGACCTGAAGAAAGGCTTGCACCTGTACGGCACGCAGGGCAACGTGGGCCTCACCAACGCCT
                                                                                         540
GCCTGTTCTAACTGTCCATACGGGTCGTTCTGGACTTCTTTCCGAACGTGGACATGCCGTGCGTCCCGTTGCACCCGGAGTGGTTGCGGA

T  D  K  I  D  R  Y  A  Q  Q  D  L  K  K  G  L  H  L  Y  G  T  Q  G  N  V  G  L  T  N  A

GGAGCATCATCCAGACCGACTTCCGCTGCTGTGGCGTCTCCAACTACACTGACTGGTTCGAGGTGTACAACGCCACGCGGGTACCTGACT
                                                                                         630
CCTCGTAGTAGGTCTGGCTGAAGGCGACGACACCGCAGAGGTTGATGTGACTGACCAAGCTCCACATGTTGCGGTGCGCCCATGGACTGA

W  S  I  I  Q  T  D  F  R  C  C  G  V  S  N  Y  T  D  W  F  E  V  Y  N  A  T  R  V  P  D

CCTGCTGCTTGGAGTTCAGTGAGAGCTGTGGGCTGCACGCCCCCGGCACCTGGTGGAAGGCGCCGTGCTACGAGACGGTGAAGGTGTGGC
                                                                                         720
GGACGACGAACCTCAAGTCACTCTCGACACCCGACGTGCGGGGGCCGTGGACCACCTTCCGCGGCACGATGCTCTGCCACTTCCACACCG

S  C  C  L  E  F  S  E  S  C  G  L  H  A  P  G  T  W  W  K  A  P  C  Y  E  T  V  K  V  W

TTCAGGAGAACCTGCTGGCTGTGGGCATCTTTGGGCTGTGCACGGCGCTGGTGCAGATCCTGGGCCTGACCTTCGCCATGACCATGTACT
                                                                                         810
AAGTCCTCTTGGACGACCGACACCCGTAGAAACCCGACACGTGCCGCGACCACGTCTAGGACCCGGACTGGAAGCGGTACTGGTACATGA

L  Q  E  N  L  L  A  V  G  I  F  G  L  C  T  A  L  V  Q  I  L  G  L  T  F  A  M  T  M  Y

GCCAAGTGGTCAAGGCAGACACCTACTGCGCGTAGGCCGCCCACCGCCGGCTTCTCTGCCAAAAGGACGCCCACGGGGAGATGGCCGCAC
                                                                                         900
CGGTTCACCAGTTCCGTCTGTGGATGACGCGCATCCGGCGGGTGGCGGCCGAAGAGACGGTTTTCCTGCGGGTGCCCCTCTACCGGCGTG

C  Q  V  V  K  A  D  T  Y  C  A  .
```

(SEQ ID NO:24)

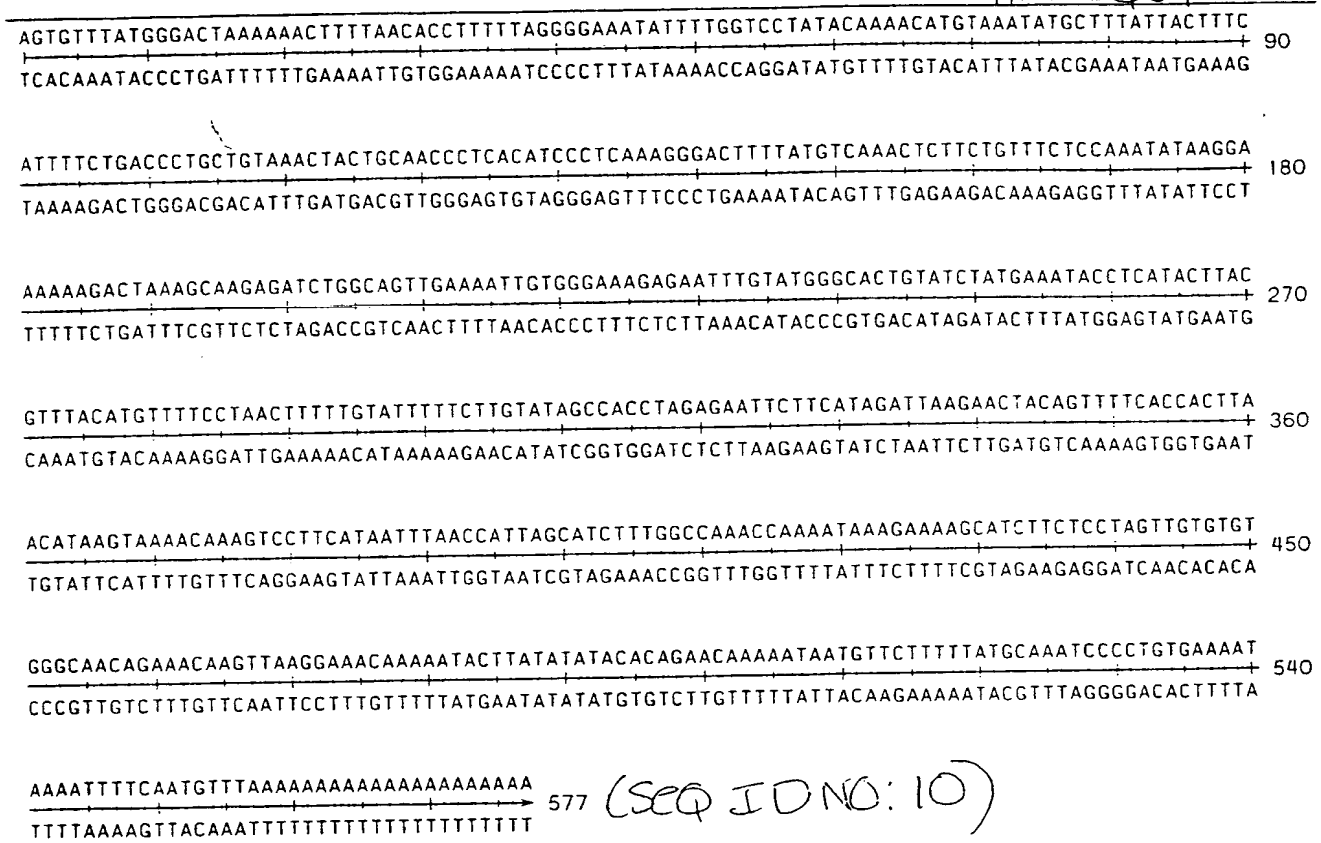
HTPBA27

```
CCACAGCTGCTTTTCCCACCACCAGCTTCGGTGTTCTGCCCCATGCTGGGAGGAGGGAGGGAGGGACAGGTGCCTGGAGCCCCCGGAACC
GGTGTCGACGAAAAGGGTGGTGGTCGAAGCCACAAGACGGGGTACGACCCTCCTCCCTCCCTCCCTGTCCACGGACCTCGGGGGCCTTGG                990

CTGTTTCTGGAAGGCCCTAGCTCAGGTGGCTTCAGGGCCTCCGGACCCCCCCTGGGAGGGGTGGCCACGTGCTGGCTGCGGAACCCAGGG
GACAAAGACCTTCCGGGATCGAGTCCACCGAAGTCCCGGAGGCCTGGGGGGGACCCTCCCCACCGGTGCACGACCGACGCCTTGGGTCCC               1080

CAGGGGTGGGAGGGGCCTCCAGCACTTTTTATATTTACGTATTCTCCAAAGCAGTGTTCACACGGGAGCCAGCCTGTGGCCCCCAGCTTC
GTCCCCACCCTCCCCGGAGGTCGTGAAAAATATAAATGCATAAGAGGTTTCGTCACAAGTGTGCCCTCGGTCGGACACCGGGGGTCGAAG               1170

CTGGAAAACAGGTTGGCGCTGGAGGAGCCGGGTCTTGGCATCCTGGAGGTGGCCCCACTGGTCCTGGTGCTCCAGGCGGGGCCGTGGACC
GACCTTTTGTCCAACCGCGACCTCCTCGGCCCAGAACCGTAGGACCTCCACCGGGGTGACCAGGACCACGAGGTCCGCCCCGGCACCTGG               1260

                                                                                    (SEQ ID NO:8)
CCTCACCTACATTCCATAGTGGGCCCGTGGGGCTCCTGGTGCATCTTAATAAAGTGTGAGCAGCAAAAAAAAAAAAAAAAAAAAAA
GGAGTGGATGTAAGGTATCACCCGGGCACCCCGAGGACCACGTAGAATTATTTCACACTCGTCGTTTTTTTTTTTTTTTTTTTTT               1345
```

HAIDQ59

```
GCGCCGCCGGGCCGCAGCATGGGGCGCTTCCGCGGGGGCCTGCGGTGCATCAAGTACCTGCTGCTTGGCTTCAACCTGCTCTTCTGGCTG
                                                                                            90
CGCGGCGGCCCGGCGTCGTACCCCGCGAAGGCGCCCCCGGACGCCACGTAGTTCATGGACGACGAACCGAAGTTGGACGAGAAGACCGAC

                  M  G  R  F  R  G  G  L  R  C  I  K  Y  L  L  L  G  F  N  L  L  F  W  L

GCTGGATCGGCCGTCATTGCTTTTGGACTATGGTTTCGGTTCGGAGGTGCCATAAAGGAGTTATCATCAGAGGACAAGTCCCCAGAGTAT
                                                                                           180
CGACCTAGCCGGCAGTAACGAAAACCTGATACCAAAGCCAAGCCTCCACGGTATTTCCTCAATAGTAGTCTCCTGTTCAGGGGTCTCATA

 A  G  S  A  V  I  A  F  G  L  W  F  R  F  G  G  A  I  K  E  L  S  S  E  D  K  S  P  E  Y

TTCTATGTGGGGCTGTATGTTCTGGTTGGAGCCGGGGCCCTGATGATGGCCGTGGGGTTCTTCGGATGCTGCGGAGCCATGCGGGAGTCG
                                                                                           270
AAGATACACCCCGACATACAAGACCAACCTCGGCCCCGGGACTACTACCGGCACCCCAAGAAGCCTACGACGCCTCGGTACGCCCTCAGC

 F  Y  V  G  L  Y  V  L  V  G  A  G  A  L  M  M  A  V  G  F  F  G  C  C  G  A  M  R  E  S

CAATGTGTGCTTGGATCATTTTTTACCTGCCTCCTGGTGATATTTGCTGCTGAAGTAACCACTGGAGTATTTGCTTTTATAGGCAAGGGG
                                                                                           360
GTTACACACGAACCTAGTAAAAAATGGACGGAGGACCACTATAAACGACGACTTCATTGGTGACCTCATAAACGAAAATATCCGTTCCCC

 Q  C  V  L  G  S  F  F  T  C  L  L  V  I  F  A  A  E  V  T  T  G  V  F  A  F  I  G  K  G

GTAGCTATCCGACATGTTCAGACCATGTATGAAGAGGCTTACAATGATTACCTTAAAGACAGGGGAAAAGGCAATGGGACACTCATCACC
                                                                                           450
CATCGATAGGCTGTACAAGTCTGGTACATACTTCTCCGAATGTTACTAATGGAATTTCTGTCCCCTTTTCCGTTACCCTGTGAGTAGTGG

 V  A  I  R  H  V  Q  T  M  Y  E  E  A  Y  N  D  Y  L  K  D  R  G  K  G  N  G  T  L  I  T

TTCCACTCAACATTTCAGTGCTGTGGAAAAGAAAGCTCCGAACAGGTCCAACCTACATGCCCAAAGGAGCTTCTAGGACACAAGAATTGC
                                                                                           540
AAGGTGAGTTGTAAAGTCACGACACCTTTTCTTTCGAGGCTTGTCCAGGTTGGATGTACGGGTTTCCTCGAAGATCCTGTGTTCTTAACG

 F  H  S  T  F  Q  C  C  G  K  E  S  S  E  Q  V  Q  P  T  C  P  K  E  L  L  G  H  K  N  C

ATCGATGAAATTGAGACCATAATCAGTGTTAAGCTCCAGCTCATTGGAATTGTCGGTATTGGAATTGCAGGTCTGACGATCTTTGGCATG
                                                                                           630
TAGCTACTTTAACTCTGGTATTAGTCACAATTCGAGGTCGAGTAACCTTAACAGCCATAACCTTAACGTCCAGACTGCTAGAAACCGTAC

 I  D  E  I  E  T  I  I  S  V  K  L  Q  L  I  G  I  V  G  I  G  I  A  G  L  T  I  F  G  M

ATATTCAGCATGGTCCTCTGCTGTGCGATACGAAACTCACGAGATGTGATATGAAGCTACTTCTACATGAAAATTGCAATCTAAAGCTTT
                                                                                           720
TATAAGTCGTACCAGGAGACGACACGCTATGCTTTGAGTGCTCTACACTATACTTCGATGAAGATGTACTTTTAACGTTAGATTTCGAAA

 I  F  S  M  V  L  C  C  A  I  R  N  S  R  D  V  I .   (SEQ ID NO:25)

CATACCAAATGTTC
              734   (SEQ ID NO:9)
GTATGGTTTACAAG
```

HAIDQ59



HHFE1540

```
TTCGGCACGAGCTGCGGGCGGTGGGCGGCTGGGCGGCCCCGGGAGCCGCGCTCTCAGTCTCTCTAGGCGCAGTCCCTTCGCCGCTTCCGG
AAGCCGTGCTCGACGCCCGCCACCCGCCGACCCGCCGGGGCCCTCGGCGCGAGAGTCAGAGAGATCCGCGTCAGGGAAGCGGCGAAGGCC

AGCCCCTGGCAGGGCCCAGAAGCCATGGCCCACTATAAGACTGAGCAGGACGACTGGCTGATCATCTACTTGAAGTATTTACTCTTTGTC
TCGGGGACCGTCCCGGGTCTTCGGTACCGGGTGATATTCTGACTCGTCCTGCTGACCGACTAGTAGATGAACTTCATAAATGAGAAACAG

                          M  A  H  Y  K  T  E  Q  D  D  W  L  I  I  Y  L  K  Y  L  L  F  V

TTCAACTTCTTCTTCTGGGTCGGGGGAGCAGCCGTCCTGGCTGTGGGCATCTGGACCCTGGTGGAGAAGAGTGGCTACCTCAGCGTCCTG
AAGTTGAAGAAGAAGACCCAGCCCCCTCGTCGGCAGGACCGACACCCGTAGACCTGGGACCACCTCTTCTCACCGATGGAGTCGCAGGAC

 F  N  F  F  F  W  V  G  G  A  A  V  L  A  V  G  I  W  T  L  V  E  K  S  G  Y  L  S  V  L

GCCTCCAGCACCTTTGCCGCCTCCGCCTACATCCTCATCTTTGCGGGCGTACTTGTCATGGTGACCGGCTTCCTGGGCTTCGGTGCCATC
CGGAGGTCGTGGAAACGGCGGAGGCGGATGTAGGAGTAGAAACGCCCGCATGAACAGTACCACTGGCCGAAGGACCCGAAGCCACGGTAG

 A  S  S  T  F  A  A  S  A  Y  I  L  I  F  A  G  V  L  V  M  V  T  G  F  L  G  F  G  A  I

CTCTGGGAGCGGAAGGGCTGCCTCTCCACGTATTTCTGCCTGTTGCTCGTCATCTTCCTGGTTGAGCTGGTGGCGGGAGTCCTGGCCCAT
GAGACCCTCGCCTTCCCGACGGAGAGGTGCATAAAGACGGACAACGAGCAGTAGAAGGACCAACTCGACCACCGCCCTCAGGACCGGGTA

 L  W  E  R  K  G  C  L  S  T  Y  F  C  L  L  L  V  I  F  L  V  E  L  V  A  G  V  L  A  H

GTGTATTACCAGAGGCTGAGTGATGAACTGAAGCAGCACTTGAACCGGACTCTGGCTGAGAACTACGGGCAGCCGGAGCACGCAGATCAC
CACATAATGGTCTCCGACTCACTACTTGACTTCGTCGTGAACTTGGCCTGAGACCGACTCTTGATGCCCGTCGGCCTCGTGCGTCTAGTG

 V  Y  Y  Q  R  L  S  D  E  L  K  Q  H  L  N  R  T  L  A  E  N  Y  G  Q  P  E  H  A  D  H

GCCTCAGTGGACCGACTCCAGCAGGATTTCAAGTGCTGCGGAAGCAACAGCTCAGCCGACTGGCAGCACAGCACGTACATCCTGTTGCGG
CGGAGTCACCTGGCTGAGGTCGTCCTAAAGTTCACGACGCCTTCGTTGTCGAGTCGGCTGACCGTCGTGTCGTGCATGTAGGACAACGCC

 A  S  V  D  R  L  Q  Q  D  F  K  C  C  G  S  N  S  S  A  D  W  Q  H  S  T  Y  I  L  L  R

GAGGCCGAGGGCCGCCAGGTGCCCGACAGCTGCTGCAAGACAGTGGTGGCGCGCTGCGGCCAGCGGGCCCACCCCTCCAACATCTATAAG
CTCCGGCTCCCGGCGGTCCACGGGCTGTCGACGACGTTCTGTCACCACCGCGCGACGCCGGTCGCCCGGGTGGGGAGGTTGTAGATATTC

 E  A  E  G  R  Q  V  P  D  S  C  C  K  T  V  V  A  R  C  G  Q  R  A  H  P  S  N  I  Y  K

GTGGAGGGAGGCTGCCTCACCAAGCTGGAGCAGTTCCTGGCCGACCACCTGCTGCTTATGGGGGCAGTGGGCATCGGGGTGGCCTGCCTG
CACCTCCCTCCGACGGAGTGGTTCGACCTCGTCAAGGACCGGCTGGTGGACGACGAATACCCCCGTCACCCGTAGCCCCACCGGACGGAC

 V  E  G  G  C  L  T  K  L  E  Q  F  L  A  D  H  L  L  L  M  G  A  V  G  I  G  V  A  C  L

CAGATCTGCGGGATGGTTCTCACCTGCTGCTTGCACCAGAGGCTCCAGCGGCATTTTTACTAATGGCAACCACCTCCTCTTCCAACTGCC
GTCTAGACGCCCTACCAAGAGTGGACGACGAACGTGGTCTCCGAGGTCGCCGTAAAAATGATTACCGTTGGTGGAGGAGAAGGTTGACGG

 Q  I  C  G  M  V  L  T  C  C  L  H  Q  R  L  Q  R  H  F  Y  .
```

(SEQ ID NO:26)

HHFEB40

CCTCAAGACAACATGTGGCACATGCCATCTGCAAGG
GGAGTTCTGTTGTACACCGTGTACGGTAGACGTTCC 936 (SEQ ID NO: 11)

HGBGV89

```
AGCTTACTTTCACTCACCGCCTGTCCTTCCTGACACCTCACCATGTGTACGGGAAAATGTGCCCGCTGTGTGGGGCTCTCCCTCATTACC 90
TCGAATGAAAGTGAGTGGCGGACAGGAAGGACTGTGGAGTGGTACACATGCCCTTTTACACGGGCGACACACCCCGAGAGGGAGTAATGG

                                         M  C  T  G  K  C  A  R  C  V  G  L  S  L  I  T

CTCTGCCTCGTCTGCATTGTGGCCAACGCCCTCCTGCTGGTACCTAATGGGGAGACCTCCTGGACCAACACCAACCATCTCAGCTTGCAA 180
GAGACGGAGCAGACGTAACACCGGTTGCGGGAGGACGACCATGGATTACCCCTCTGGAGGACCTGGTTGTGGTTGGTAGAGTCGAACGTT

 L  C  L  V  C  I  V  A  N  A  L  L  L  V  P  N  G  E  T  S  W  T  N  T  N  H  L  S  L  Q

GTCTGGCTCATGGGCGGCTTCATTGGCGGGGGCCTAATGGTACTGTGTCCAGGGATTGCAGCCGTTCGGGCAGGGGGCAAGGGCTGCTGT 270
CAGACCGAGTACCCGCCGAAGTAACCGCCCCCGGATTACCATGACACAGGTCCCTAACGTCGGCAAGCCCGTCCCCCGTTCCCGACGACA

 V  W  L  M  G  G  F  I  G  G  G  L  M  V  L  C  P  G  I  A  A  V  R  A  G  G  K  G  C  C

GGTGCTGGGTGCTGTGGAAACCGCTGCAGGATGCTGCGCTCGGTCTTCTCCTCGGCGTTCGGGGTGCTTGGTGCCATCTACTG
CCACGACCCACGACACCTTTGGCGACGTCCTACGACGCGAGCCAGAAGAGGAGCCGCAAGCCCCACGAACCACGGTAGATGACGG

 G  A  G  C  C  G  N  R  C  R  M  L  R  S  V  F  S  S  A  F  G  V  L  G  A  I  Y  C  L

GTGTCTGGAGCTGGGCTCCGAAATGGACCCAGATGCTTAATGAACGGCGAGTGGGGCTACCACTTCGAAGACACCGCGGGAGCTTACTTG 450
CACAGACCTCGACCCGAGGCTTTACCTGGGTCTACGAATTACTTGCCGCTCACCCCGATGGTGAAGCTTCTGTGGCGCCCTCGAATGAAC

 V  S  G  A  G  L  R  N  G  P  R  C  L  M  N  G  E  W  G  Y  H  F  E  D  T  A  G  A  Y  L

CTCAACCGCACTCTATGGGATCGGTGCGAGGCGCCCCCTCGCGTGGTCCCCTGGAATGTGACGCTCTTCTCGCTGCTGGTGGCCGCCTCC 540
GAGTTGGCGTGAGATACCCTAGCCACGCTCCGCGGGGGAGCGCACCAGGGGACCTTACACTGCGAGAAGAGCGACGACCACCGGCGGAGG

 L  N  R  T  L  W  D  R  C  E  A  P  P  R  V  V  P  W  N  V  T  L  F  S  L  L  V  A  A  S

TGCCTGGAGATAGTACTGTGTGGGATCCAGCTGGTGAACGCGACCATTGGTGTCTTCTGCGGCGATTGCAGGAAAAAACAGGACACACCT 630
ACGGACCTCTATCATGACACACCCTAGGTCGACCACTTGCGCTGGTAACCACAGAAGACGCCGCTAACGTCCTTTTTTGTCCTGTGTGGA

 C  L  E  I  V  L  C  G  I  Q  L  V  N  A  T  I  G  V  F  C  G  D  C  R  K  K  Q  D  T  P

CACTGAGGCTCCACTGACCGCCGGGTTACACCTGCTCCTTCCTGGACGCTCACTCCCTTGCTCGCTAGAATAAACTGCTTTGCGCTCTCA 720
GTGACTCCGAGGTGACTGGCGGCCCAATGTGGACGAGGAAGGACCTGCGAGTGAGGGAACGAGCGATCTTATTTGACGAAACGCGAGAGT

 H . (SEQ ID NO:27)

AAAAAAAAAAAAAAAAAC 738 (SEQ ID NO:12)
TTTTTTTTTTTTTTTTTG
```

HUVB880

```
GGCACGAGAGATTGTCGGCTGCGGGTATATTCCAATTCCCCGTCTCCTCATGAATATGAAGTGAAGGGCTCTGACCCTGGAAGTGGTTCT
CCGTGCTCTCTAACAGCCGACGCCCATATAAGGTTAAGGGGCAGAGGAGTACTTATACTTCACTTCCCGAGACTGGGACCTTCACCAAGA

AAGCAGGGCAAAATGGGGTCTCGGAAGTGTGGAGGCTGCCTAAGTTGTTTGCTGATTCCGCTTGCACTTTGGAGTATAATCGTGAACATA
TTCGTCCCGTTTTACCCCAGAGCCTTCACACCTCCGACGGATTCAACAAACGACTAAGGCGAACGTGAAACCTCATATTAGCACTTGTAT

                M  G  S  R  K  C  G  G  C  L  S  C  L  L  I  P  L  A  L  W  S  I  I  V  N  I

TTATTGTATTTCCCGAATGGGCAAACTTCCTATGCATCCAGCAATAAACTCACCAACTACGTGTGGTATTTTGAAGGAATCTGTTTCTCA
AATAACATAAAGGGCTTACCCGTTTGAAGGATACGTAGGTCGTTATTTGAGTGGTTGATGCACACCATAAAACTTCCTTAGACAAAGAGT

 L  L  Y  F  P  N  G  Q  T  S  Y  A  S  S  N  K  L  T  N  Y  V  W  Y  F  E  G  I  C  F  S

GGCATCATGATGCTTATAGTAACAACAGTTCTTCTGGTACTGGAGAATAATAACAACTATAAATGTTGCCAGAGTGAAAACTGCAGCAAA
CCGTAGTACTACGAATATCATTGTTGTCAAGAAGACCATGACCTCTTATTATTGTTGATATTTACAACGGTCTCACTTTTGACGTCGTTT

 G  I  M  M  L  I  V  T  T  V  L  L  V  L  E  N  N  N  N  Y  K  C  C  Q  S  E  N  C  S  K

AAATATGTGACACTGCTGTCAATTATCTTTTCTTCCCTCGGAATTGCTTTTTCTGGATACTGCCTGGTCATCTCTGCCTTGGGTCTTGTC
TTTATACACTGTGACGACAGTTAATAGAAAAGAAGGGAGCCTTAACGAAAAAGACCTATGACGGACCAGTAGAGACGGAACCCAGAACAG

 K  Y  V  T  L  L  S  I  I  F  S  S  L  G  I  A  F  S  G  Y  C  L  V  I  S  A  L  G  L  V

CAAGGGCCATATTGCCGCACCCTTGATGGCTGGGAGTATGCTTTTGAAGGCACTGCTGGACGTTTCCTTACAGATTCTAGCATATGGATT
GTTCCCGGTATAACGGCGTGGGAACTACCGACCCTCATACGAAAACTTCCGTGACGACCTGCAAAGGAATGTCTAAGATCGTATACCTAA

 Q  G  P  Y  C  R  T  L  D  G  W  E  Y  A  F  E  G  T  A  G  R  F  L  T  D  S  S  I  W  I

CAGTGCCTGGAACCTGCACATGTTGTGGAGTGGAACATCATTTTATTTTCCATTCTCATAACCCTCAGTGGGCTTCAAGTGATCATCTGC
GTCACGGACCTTGGACGTGTACAACACCTCACCTTGTAGTAAAATAAAAGGTAAGAGTATTGGGAGTCACCCGAAGTTCACTAGTAGACG

 Q  C  L  E  P  A  H  V  V  E  W  N  I  I  L  F  S  I  L  I  T  L  S  G  L  Q  V  I  I  C

CTCATCAGAGTAGTCATGCAACTATCCAAGATACTGTGTGGAAGCTATTCAGTGATCTTCCAGCCTGGAATCATTTGAATAAGGACAAAA
GAGTAGTCTCATCAGTACGTTGATAGGTTCTATGACACACCTTCGATAAGTCACTAGAAGGTCGGACCTTAGTAAACTTATTCCTGTTTT

 L  I  R  V  V  M  Q  L  S  K  I  L  C  G  S  Y  S  V  I  F  Q  P  G  I  I .  (SEQ ID NO:28)

TGTTTTCCATTATCAAGACATGGCCATCTATCTAAATATTATATCAACTGTGTTAGACTTGAGGGCAATATTGAAAATGATGGTGCTTTC
ACAAAAGGTAATAGTTCTGTACCGGTAGATAGATTTATAATATAGTTGACACAATCTGAACTCCCGTTATAACTTTTACTACCACGAAAG

TGCATTTGGTGTTTATTTGTAAAAAATTTGCAGTCCTCACTGCACATGCAAGTATACCACCCTTCCATTTAGTATGTTTTTTAAGTAATA
ACGTAAACCACAAATAAACATTTTTTAAACGTCAGGAGTGACGTGTACGTTCATATGGTGGGAAGGTAAATCATACAAAAAATTCATTAT
```

HUVBB80

TGCATCAGAAACTTCAGAAATACTTCTGCCCTTTGATCAAACAAATCCATTTCCAAGAATCTGTACTAGGGAAGTAAATAAGAATATGAG 990
ACGTAGTCTTTGAAGTCTTTATGAAGACGGGAAACTAGTTTGTTTAGGTAAAGGTTCTTAGACATGATCCCTTCATTTATTCTTATACTC

(SEQ ID NO:B

AGAAACCTTTATGCAATATGTATATTGCAACATTATTTAATATTCTGGAAAATTGGAAACACCCCAAAATTCTAACTCAAA 1071
TCTTTGGAAATACGTTATACATATAACGTTGTAATAAATTATAAGACCTTTTAACCTTTGTGGGGTTTTAAGATTGAGTTT

HJACE54

```
TTTGTGGAGGGCAGCAGAGAGTACCCAGCTGGACATCCTTTCCTGCTGATGAGCCCCAGGCTGGAGGTGCCCTGCTCACATGCTCTTCCC
+ 90
AAACACCTCCCGTCGTCTCTCATGGGTCGACCTGTAGGAAAGGACGACTACTCGGGGTCCGACCTCCACGGGACGAGTGTACGAGAAGGG

                                 M  S  P  R  L  E  V  P  C  S  H  A  L  P

CAGGGTCTCTCGCCTGGGCAGGTCATCATAGTACGGGGACTGGTCTTGCAAGAGCCGAAGCATTTTACTGTGAGCCTGAGGGACCAGGCT
+ 180
GTCCCAGAGAGCGGACCCGTCCAGTAGTATCATGCCCCTGACCAGAACGTTCTCGGCTTCGTAAAATGACACTCGGACTCCCTGGTCCGA

 Q  G  L  S  P  G  Q  V  I  I  V  R  G  L  V  L  Q  E  P  K  H  F  T  V  S  L  R  D  Q  A

GCCCATGCTCCTGTGACACTCAGGGCCTCCTTCGCAGACAGAACTCTGGCCTGGATCTCCCGCTGGGGGCAGAAGAAACTGATCTCAGCC
+ 270
CGGGTACGAGGACACTGTGAGTCCCGGAGGAAGCGTCTGTCTTGAGACCGGACCTAGAGGGCGACCCCCGTCTTCTTTGACTAGAGTCGG

 A  H  A  P  V  T  L  R  A  S  F  A  D  R  T  L  A  W  I  S  R  W  G  Q  K  K  L  I  S  A

CCCTTCCTCTTTTACCCCCAGAGATTCTTTGAGGTGCTGCTCCTGTTCCAGGAGGGAGGGCTGAAGCTGGCGCTCAATGGGCAGGGGCTG
+ 360
GGGAAGGAGAAAATGGGGGTCTCTAAGAAACTCCACGACGAGGACAAGGTCCTCCCTCCCGACTTCGACCGCGAGTTACCCGTCCCCGAC

 P  F  L  F  Y  P  Q  R  F  F  E  V  L  L  L  F  Q  E  G  G  L  K  L  A  L  N  G  Q  G  L

GGGGCCACCAGCATGAACCAGCAGGCCCTGGAGCAGCTGCGGGAGCTCCGGATCAGTGGAAGTGTCCAGCTCTACTGTGTCCACTCCTGA
+ 450
CCCCGGTGGTCGTACTTGGTCGTCCGGGACCTCGTCGACGCCCTCGAGGCCTAGTCACCTTCACAGGTCGAGATGACACAGGTGAGGACT

 G  A  T  S  M  N  Q  Q  A  L  E  Q  L  R  E  L  R  I  S  G  S  V  Q  L  Y  C  V  H  S  .   (SEQ ID NO:29)

AGGATGGTTCCAGGAAATACCGCAGAAAACAAGAGTCAGCCACTCCCCAGGGCCCCACTCTCCTCCCCTCATTAAACCATCCACCTGAAC
+ 540
TCCTACCAAGGTCCTTTATGGCGTCTTTTGTTCTCAGTCGGTGAGGGGTCCCGGGGTGAGAGGAGGGGAGTAATTTGGTAGGTGGACTTG

ACCAGCACATCAGGGCCTGGTTCACCTCTGGGGTCACGAGACTGAGTCTACAGGAGCTTTGGGCCTGAGGGAAGGCACAAGAGTGCAAAG
+ 630
TGGTCGTGTAGTCCCGGACCAAGTGGAGACCCCAGTGCTCTGACTCAGATGTCCTCGAAACCCGGACTCCCTTCCGTGTTCTCACGTTTC

GTTCCTCGAACTCTGCACCTTCCTCCACCAGGAGCCTGGGATATGGCTCCATCTGCCTTCAGGGCCTGGACTGCACTCACAGAGGCAAGT
+ 720
CAAGGAGCTTGAGACGTGGAAGGAGGTGGTCCTCGGACCCTATACCGAGGTAGACGGAAGTCCCGGACCTGACGTGAGTGTCTCCGTTCA

GTTGTAGACTAACAAAGATACTCCAAAATACAATGGCTTAAAGAATGTGGTCATTTATTCTTTATTATTTATTTATTTGTGGTCAAATAA
+ 810
CAACATCTGATTGTTTCTATGAGGTTTTATGTTACCGAATTTCTTACACCAGTAAATAAGAAATAATAAATAAATAAACACCAGTTTATT

ATAAATAAGGTTATTTATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
> 865 (SEQ ID NO: 14)
TATTTATTCCAATAAATAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
```

HROAD63


```
GCACGAGAGACGACATCAGAGATGAGGACAGCATTGCTGCTCCTTGCAGCCCTGGCTGTGGCTACAGGGCCAGCCCTTACCCTGCGCTGC
                                                                                          90
CGTGCTCTCTGCTGTAGTCTCTACTCCTGTCGTAACGACGAGGAACGTCGGGACCGACACCGATGTCCCGGTCGGGAATGGGACGCGACG

                  M  R  T  A  L  L  L  L  A  A  L  A  V  A  T  G  P  A  L  T  L  R  C

CACGTGTGCACCAGCTCCAGCAACTGCAAGCATTCTGTGGTCTGCCCGGCCAGCTCTCGCTTCTGCAAGACCACGAACACAGTGGAGCCT
                                                                                         180
GTGCACACGTGGTCGAGGTCGTTGACGTTCGTAAGACACCAGACGGGCCGGTCGAGAGCGAAGACGTTCTGGTGCTTGTGTCACCTCGGA

 H  V  C  T  S  S  S  N  C  K  H  S  V  V  C  P  A  S  S  R  F  C  K  T  T  N  T  V  E  P

CTGAGGGCTTCCCCGAAAGTCTGGGACCAGGTCCAGGTGGGCATGGAATGCTGATGACTTGGAGCAGGCCCCACAGACCCCACAGAGGAT
                                                                                         270
GACTCCCGAAGGGGCTTTCAGACCCTGGTCCAGGTCCACCCGTACCTTACGACTACTGAACCTCGTCCGGGGTGTCTGGGGTGTCTCCTA

 L  R  A  S  P  K  V  W  D  Q  V  Q  V  G  M  E  C  .

GAAGCCACCCCACAGAGGATGCAGCCCCCAGCTGCATGGAAGGTGGAGGACAGAAGCCCTGTGGATCCCCGGATTTCACACTCCTTCTGT
                                                                                         360
CTTCGGTGGGGTGTCTCCTACGTCGGGGGTCGACGTACCTTCCACCTCCTGTCTTCGGGACACCTAGGGGCCTAAAGTGTGAGGAAGACA

TTTGTTGCCGTTTATTTTTGTACTCAAATCTCTACATGGAGATAAATGATTTAAACCAGTAAAAAAAAAAAAAAAAAAAAA
                                                                                         441
AAACAACGGCAAATAAAAACATGAGTTTAGAGATGTACCTCTATTTACTAAATTTGGTCATTTTTTTTTTTTTTTTTTTTT
```

(SEQ ID NO: 30)

(SEQ ID NO: 15)

```
AGCGGGCCCGAACCCTCGTGTGAAGGGTGCAGTACCTAAGCCGGAGCGGGGTAGAGGCGGGCCGGCACCCCCTTCTGACCTCCAGTGCCG
TCGCCCGGGCTTGGGAGCACACTTCCCACGTCATGGATTCGGCCTCGCCCCATCTCCGCCCGGCCGTGGGGGAAGACTGGAGGTCACGGC  90

CCGGCCTCAAGATCAGACATGGCCCAGAACTTGAAGGACTTGGCGGGACGGCTGCCCGCCGGGCCCCGGGGCATGGGCACGGCCCTGAAG
GGCCGGAGTTCTAGTCTGTACCGGGTCTTGAACTTCCTGAACCGCCCTGCCGACGGGCGGCCCGGGGCCCCGTACCCGTGCCGGGACTTC 180

                  M  A  Q  N  L  K  D  L  A  G  R  L  P  A  G  P  R  G  M  G  T  A  L  K

CTGTTGCTGGGGGCCGGCGCCGTGGCCTACGGTGTGCGCGAATCTGTGTTCACCGTGGAAGGCGGGCACAGAGCCATCTTCTTCAATCGG
GACAACGACCCCCGGCCGCGGCACCGGATGCCACACGCGCTTAGACACAAGTGGCACCTTCCGCCCGTGTCTCGGTAGAAGAAGTTAGCC 270

 L  L  L  G  A  G  A  V  A  Y  G  V  R  E  S  V  F  T  V  E  G  G  H  R  A  I  F  F  N  R

ATCGGTGGAGTGCAGCAGGACACTATCCTGGCCGAGGGCCTTCACTTCAGGATCCCTTGGTTCCAGTACCCCATTATCTATGACATTCGG
TAGCCACCTCACGTCGTCCTGTGATAGGACCGGCTCCCGGAAGTGAAGTCCTAGGGAACCAAGGTCATGGGGTAATAGATACTGTAAGCC 360

 I  G  G  V  Q  Q  D  T  I  L  A  E  G  L  H  F  R  I  P  W  F  Q  Y  P  I  I  Y  D  I  R

GCCAGACCTCGAAAAATCTCCTCCCCTACAGGCTCCAAAGACCTACAGATGGTGAATATCTCCCTGCGAGTGTTGTCTCGACCCAATGCT
CGGTCTGGAGCTTTTTAGAGGAGGGGATGTCCGAGGTTTCTGGATGTCTACCACTTATAGAGGGACGCTCACAACAGAGCTGGGTTACGA 450

 A  R  P  R  K  I  S  S  P  T  G  S  K  D  L  Q  M  V  N  I  S  L  R  V  L  S  R  P  N  A

CAGGAGCTTCCTAGCATGTACCAGCGCCTAGGGCTGGACTACGAGGAACGAGTGTTGCCGTCCATTGTCAACGAGGTGCTCAAGAGTGTG
GTCCTCGAAGGATCGTACATGGTCGCGGATCCCGACCTGATGCTCCTTGCTCACAACGGCAGGTAACAGTTGCTCCACGAGTTCTCACAC 540

 Q  E  L  P  S  M  Y  Q  R  L  G  L  D  Y  E  E  R  V  L  P  S  I  V  N  E  V  L  K  S  V

GTGGCCAAGTTCAATGCCTCACAGCTGATCACCCAGCGGGCCCAGGTATCCCTGTTGATCCGCCGGGAGCTGACAGAGAGGGCCAAGGAC
CACCGGTTCAAGTTACGGAGTGTCGACTAGTGGGTCGCCCGGGTCCATAGGGACAACTAGGCGGCCCTCGACTGTCTCTCCCGGTTCCTG 630

 V  A  K  F  N  A  S  Q  L  I  T  Q  R  A  Q  V  S  L  L  I  R  R  E  L  T  E  R  A  K  D

TTCAGCCTCATCCTGGATGATGTGGCCATCACAGAGCTGAGCTTTAGCCGAGAGTACACAGCTGCTGTAGAAGCCAAACAAGTGGCCCAG
AAGTCGGAGTAGGACCTACTACACCGGTAGTGTCTCGACTCGAAATCGGCTCTCATGTGTCGACGACATCTTCGGTTTGTTCACCGGGTC 720

 F  S  L  I  L  D  D  V  A  I  T  E  L  S  F  S  R  E  Y  T  A  A  V  E  A  K  Q  V  A  Q

CAGGAGGCCCAGCGGGCCCAATTCTTGGTAGAAAAAGCAAAGCAGGAACAGCGGCAGAAAATTGTGCAGGCCGAGGGTGAGGCCGAGGCT
GTCCTCCGGGTCGCCCGGGTTAAGAACCATCTTTTTCGTTTCGTCCTTGTCGCCGTCTTTTAACACGTCCGGCTCCCACTCCGGCTCCGA 810

 Q  E  A  Q  R  A  Q  F  L  V  E  K  A  K  Q  E  Q  R  Q  K  I  V  Q  A  E  G  E  A  E  A

GCCAAGATGCTTGGAGAAGCACTGAGCAAGAACCCTGGCTACATCAAACTTCGCAAGATTCGAGCAGCCCAGAATATCTCCAAGACGATC
CGGTTCTACGAACCTCTTCGTGACTCGTTCTTGGGACCGATGTAGTTTGAAGCGTTCTAAGCTCGTCGGGTCTTATAGAGGTTCTGCTAG 900

 A  K  M  L  G  E  A  L  S  K  N  P  G  Y  I  K  L  R  K  I  R  A  A  Q  N  I  S  K  T  I
```

HM1046546

GCCACATCACAGAATCGTATCTATCTCACAGCTGACAACCTTGTGCTGAACCTACAGGATGAAAGTTTCACCAGGGGAAGTGACAGCCTC 990

CGGTGTAGTGTCTTAGCATAGATAGAGTGTCGACTGTTGGAACACGACTTGGATGTCCTACTTTCAAAGTGGTCCCCTTCACTGTCGGAG

A T S Q N R I Y L T A D N L V L N L Q D E S F T R G S D S L .

ATCAAGGGTAAGAAATGAGCCTAGTCACCAAGAACTCCACCCCCACAAGAAGTGGATCTGCTTCTCCAGTTTTTGA 1068 (SEQ ID NO: 16

TAGTTCCCATTCTTTACTCGGATCAGTGGTTCTTGAGGTGGGGGTGTTCTTCACCTAGACGAAGAGGTCAAAAACT

I K G K K . (SEQ ID NO: 31)

```
GGCACGAGATGACATCACTAAGTGGCCGATCTGCACAGAGCAGGCCAGGAGCAACCACACAGGCTTCCTGCACATGGACTGCGAGATCAA  90
CCGTGCTCTACTGTAGTGATTCACCGGCTAGACGTGTCTCGTCCGGTCCTCGTTGGTGTGTCCGAAGGACGTGTACCTGACGCTCTAGTT

                                                                          M  D  C  E  I  K

GGGCCGCCCCTGCTGCATCGGCACCAAGGGCAGCTGTGAGATCACCACCCGGGAATACTGTGAGTTCATGCACGGCTATTTCCATGAGGA  180
CCCGGCGGGGACGACGTAGCCGTGGTTCCCGTCGACACTCTAGTGGTGGGCCCTTATGACACTCAAGTACGTGCCGATAAAGGTACTCCT

  G  R  P  C  C  I  G  T  K  G  S  C  E  I  T  T  R  E  Y  C  E  F  M  H  G  Y  F  H  E  E

AGCAACACTCTGCTCCCAGGTGAGGCGAGGCAGGCCTGGAGTAGTGGAGGAGAGGACGCTGGGCATGGCAGCCTGCTGGGGCCGGGGCTC  270
TCGTTGTGAGACGAGGGTCCACTCCGCTCCGTCCGGACCTCATCACCTCCTCTCCTGCGACCCGTACCGTCGGACGACCCCGGCCCCGAG

  A  T  L  C  S  Q  V  R  R  G  R  P  G  V  V  E  E  R  T  L  G  M  A  A  C  W  G  R  G  S

ACGCACTCCCTCCCATGTCGGAGCCTCAGACTCAGGCTGCTTCTGGGGCGCTGAGCACCATATGCCCATTCCCAGGTGCACTGTTTTGGA  360
TGCGTGAGGGAGGGTACAGCCTCGGAGTCTGAGTCCGACGAAGACCCCGCGACTCGTGGTATACGGGTAAGGGTCCACGTGACAAAACCT

  R  T  P  S  H  V  G  A  S  D  S  G  C  F  W  G  A  E  H  H  M  P  I  P  R  C  T  V  L  D

CAAGGTGTGTTGGGCTGCTGCCTTCCTCAACCCTGAGGTCCCAGATCAGTTTTACAGGTCTGGCTGTCTCTTTTCCTACATGTTGGGTAA  450
GTTCCACACAACCCGACGACGGAAGGAGTTGGGACTCCAGGGTCTAGTCAAAATGTCCAGACCGACAGAGAAAAGGATGTACAACCCATT

  K  V  C  W  A  A  A  F  L  N  P  E  V  P  D  Q  F  Y  R  S  G  C  L  F  S  Y  M  L  G  K

GAGGTCCTCAATGCCCCCGAACCCGACCCCTGTGATGGACACCCAGGCGGACCCCTGGGGAAAGGTTCCTGGGCCAGGGTATGGTCGGTC  540
CTCCAGGAGTTACGGGGGCTTGGGCTGGGGACACTACCTGTGGGTCCGCCTGGGGACCCCTTTCCAAGGACCCGGTCCCATACCAGCCAG

  R  S  S  M  P  P  N  P  T  P  V  M  D  T  Q  A  D  P  W  G  K  V  P  G  P  G  Y  G  R  S

CAACCTGCCGAAGACTACTGCTCCTGAAGTGTCTGGATGAAGGCCGCTGCCTGGTGTGTCCCTCCCCCAGTGTGGGTGCACTGCCCTCGG  630
GTTGGACGGCTTCTGATGACGAGGACTTCACAGACCTACTTCCGGCGACGGACCACACAGGGAGGGGGTCACACCCACGTGACGGGAGCC

  N  L  P  K  T  T  A  P  E  V  S  G  .  (SEQ ID NO:32)

TGTCTTGTGGGTCTTTTCAAATGACATCCCTGAAGGGGACCTGGAGGAAGGTGGTCCGGCTGGCACCCTATCGC  704  (SEQ ID NO:17
ACAGAACACCCAGAAAAGTTTACTGTAGGGACTTCCCCTGGACCTCCTTCCACCAGGCCGACCGTGGGATAGCG
```

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 13, ~~line~~ Table 1

B. IDENTIFICATION OF DEPOSIT — Further deposits are identified on an additional sheet ☐

Name of depositary institution
American Type Culture Collection

Address of depositary institution *(including postal code and country)*
12301 Parklawn Drive
Rockville, Maryland 20852
United States of America

Date of deposit	Accession Number
May 16, 1997	209053

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* — This information is continued on an additional sheet ☐

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☑ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer Yvette Simms Paralegal Specialist IAPD-PCT Operations (703) 305-9746	Authorized officer

Form PCT/RO/134 (July 1992)

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 13, line Table 1.

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet ☐

Name of depositary institution

American Type Culture Collection

Address of depositary institution *(including postal code and country)*

12301 Parklawn Drive
Rockville, Maryland 20852
United States of America

Date of deposit	Accession Number
May 16, 1997	209053

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet ☐

EUROPE

In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused of withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☑ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer Yvette Simms Parajegal Specialist IAPD-PCT Operations	Authorized officer

Form PCT/RO/134 (July 1992)

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 13 line/ Table 1

B. IDENTIFICATION OF DEPOSIT — Further deposits are identified on an additional sheet ☐

Name of depositary institution
American Type Culture Collection

Address of depositary institution *(including postal code and country)*
12301 Parklawn Drive
Rockville, Maryland 20852
United States of America

Date of deposit	Accession Number
May 16, 1997	209054

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* — This information is continued on an additional sheet ☐

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☑ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer Yvette Gimms Paralegal Specialist APD-PCT Operations	Authorized officer

Form PCT/RO/134 (July 1992)

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

REC'D 23 FEB 1998
WIPO PCT

A. The indications made below relate to the microorganism referred to in the description on page 13, ~~line~~ Table 1

B. IDENTIFICATION OF DEPOSIT — Further deposits are identified on an additional sheet ☐

Name of depositary institution

American Type Culture Collection

Address of depositary institution *(including postal code and country)*

12301 Parklawn Drive
Rockville, Maryland 20852
United States of America

Date of deposit	Accession Number
May 16, 1997	209054

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* — This information is continued on an additional sheet ☐

EUROPE

In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only

☑ This sheet was received with the international application

Authorized officer

Yvette Simms
Paralegal Specialist
IAPD-PCT Operations
(703) 305-3746

For International Bureau use only

☐ This sheet was received by the International Bureau on:

Authorized officer

What is claimed is:

1. An isolated polynucleotide comprising a nucleotide sequence that has at least 80% identity over its entire length to a nucleotide sequence encoding the receptor polypeptide of SEQ ID NO:Y; or a nucleotide sequence complementary to said isolated polynucleotide.

2. The polynucleotide of claim 1 wherein said polynucleotide comprises the nucleotide sequence contained in SEQ ID NO:X encoding the receptor polypeptide of SEQ ID NO:Y.

3. The polynucleotide of claim 1 wherein said polynucleotide comprises a nucleotide sequence that is at least 80% identical to that of SEQ ID NO:X over its entire length.

4. The polynucleotide of claim 3 which is polynucleotide of SEQ ID NO: X.

5. The polynucleotide of claim 1 which is DNA or RNA.

6. A DNA or RNA molecule comprising an expression system, wherein said expression system is capable of producing a receptor polypeptide comprising an amino acid sequence, which has at least 80% identity with the polypeptide of SEQ ID NO:Y when said expression system is present in a compatible host cell.

7. A host cell comprising the expression system of claim 6.

8. A process for producing a receptor polypeptide comprising culturing a host of claim 7 under conditions sufficient for the production of said polypeptide and recovering the polypeptide from the culture.

9. A process for producing a cell which produces a receptor polypeptide thereof comprising transforming or transfecting a host cell with the expression system of claim 6 such that the host cell, under appropriate culture conditions, produces a receptor polypeptide.

10. A receptor polypeptide comprising an amino acid sequence which is at least 80% identical to the amino acid sequence of SEQ ID NO:Y over its entire length.

11. The polypeptide of claim 10 which comprises the amino acid sequence of SEQ ID NO:Y.

12. An antibody immunospecific for the receptor polypeptide of claim 10.

13. A method for the treatment of a subject in need of enhanced activity or expression of receptor polypeptide of claim 10 comprising:

(a) administering to the subject a therapeutically effective amount of an agonist to said receptor; and/or

(b) providing to the subject polynucleotide of claim 1 in a form so as to effect production of said receptor activity in vivo.

14. A method for the treatment of a subject having need to inhibit activity or expression of the receptor polypeptide of claim 10 comprising:

(a) administering to the subject a therapeutically effective amount of an antagonist to said receptor; and/or

(b) administering to the subject a nucleic acid molecule that inhibits the expression of the nucleotide sequence encoding said receptor; and/or

(c) administering to the subject a therapeutically effective amount of a polypeptide that competes with said receptor for its ligand.

15. A process for diagnosing a disease or a susceptibility to a disease in a subject related to expression or activity of the receptor polypeptide of claim 10 in a subject comprising:

(a) determining the presence or absence of a mutation in the nucleotide sequence encoding said receptor polypeptide in the genome of said subject; and/or

(b) analyzing for the presence or amount of the receptor polypeptide expression in a sample derived from said subject.

16. A method for identifying agonists to the receptor polypeptide of claim 10 comprising:

(a) contacting cells produced by claim 9 with a candidate compound; and

(b) determining whether the candidate compound effects a signal generated by activation of the receptor polypeptide.

17. An agonist identified by the method of claim 16.

18. The method for identifying antagonists to the receptor polypeptide of claim 10 comprising:

(a) contacting said cell produced by claim 9 with an agonist; and

(b) determining whether the signal generated by said agonist is diminished in the presence of a candidate compound.

19. An antagonist identified by the method of claim 18.

20. An isolated receptor polynucleotide comprising a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence having at least 80% identity to a nucleotide sequence encoding the receptor polypeptide expressed by the cDNA insert deposited at the ATCC; and

(b) a nucleotide sequence complementary to the nucleotide sequence of (a).

21. A recombinant host cell produced by a method of Claim 9 or a membrane thereof expressing a receptor polypeptide.

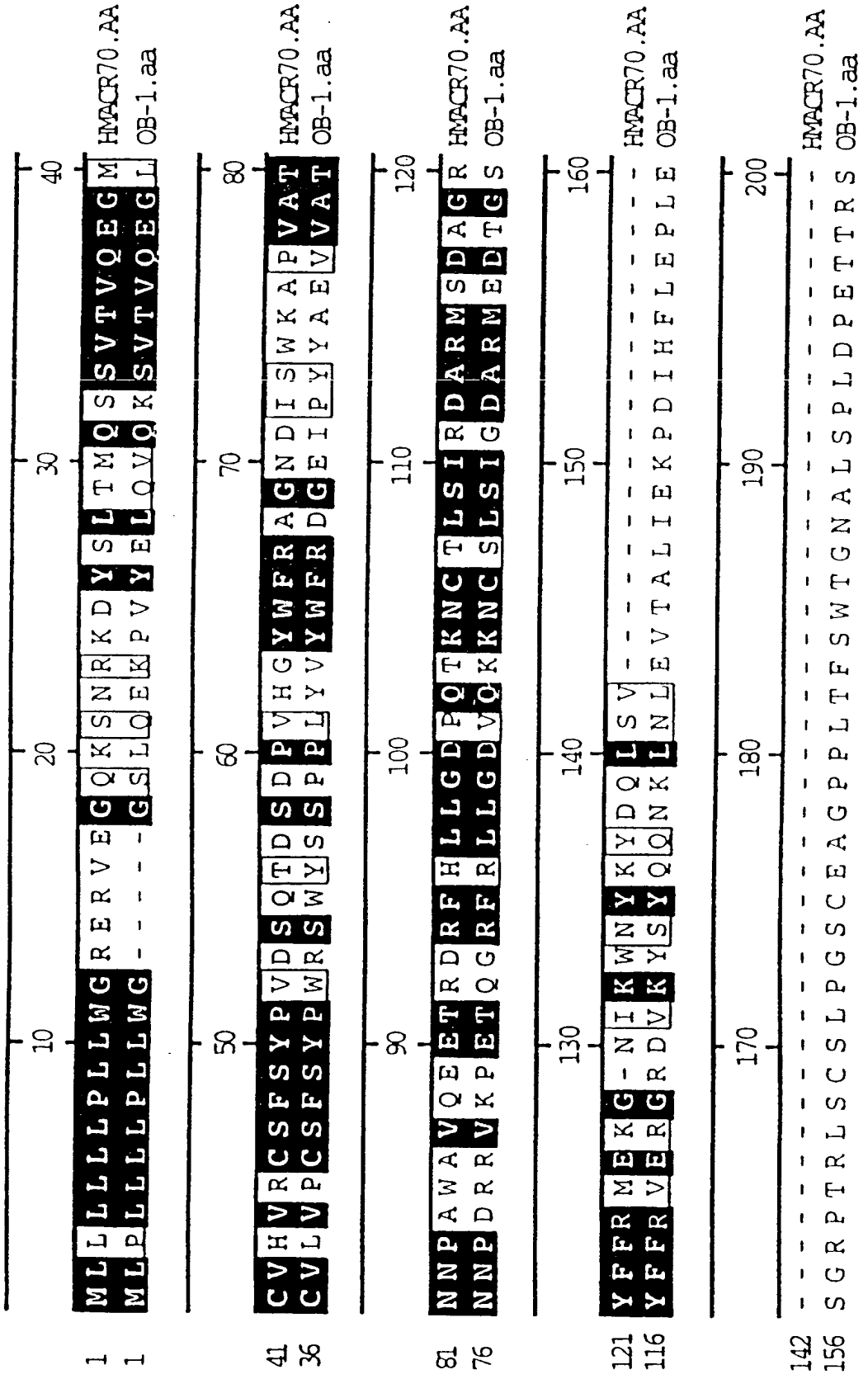


FIG.1A

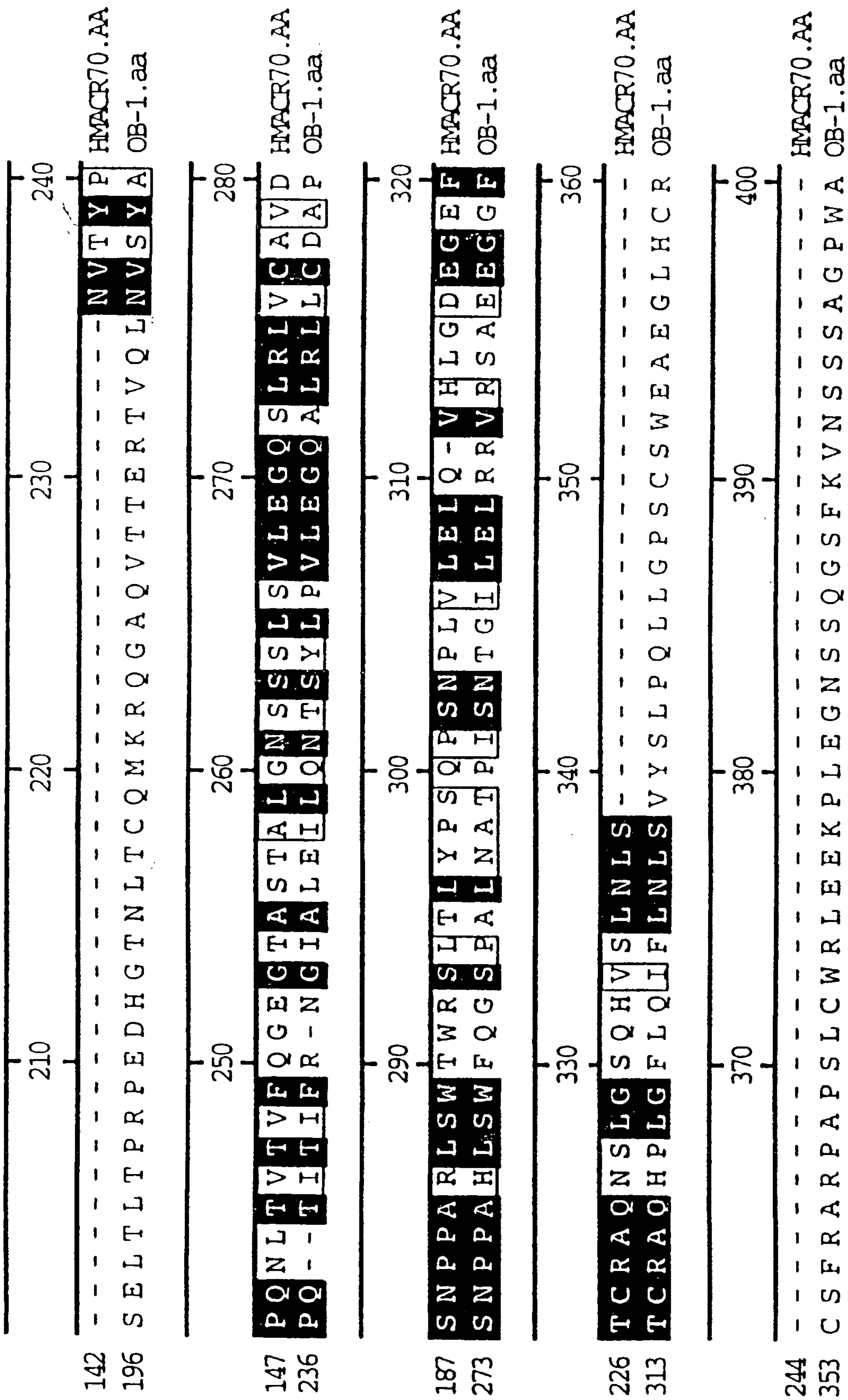

210 220 230 240
142 - N V T Y P HMACR70.AA
196 S E L T L T P R P E D H G T N L T C Q M K R Q G A Q V T T E R T V Q L N V S Y A OB-1.aa
250 260 270 280
147 P Q N L T V T V F Q G E G T A S T A L G N S S S L S V L E G Q S L R L V C A V D HMACR70.AA
236 P Q - T I T I F R - N G I A L E I L Q N T S Y L P V L E G Q A L R L L C D A P OB-1.aa
290 300 310 320
187 S N P P A R L S W T W R S L T L Y P S Q P S N P L V L E L Q - V H L G D E G E F HMACR70.AA
273 S N P P A H L S W F Q G S P A L N A T P I S N T G I L E L R R V R S A E E G F OB-1.aa
330 340 350 360
226 T C R A Q N S L G S Q H V S L N L S - HMACR70.AA
313 T C R A Q H P L G F L Q I F L N L S V Y S L P Q L L G P S C S W E A E G L H C R OB-1.aa
370 380 390 400
244 - HMACR70.AA
353 C S F R A R P A P S L C W R L E E K P L E G N S S Q G S F K V N S S S A G P W A OB-1.aa

FIG.1B

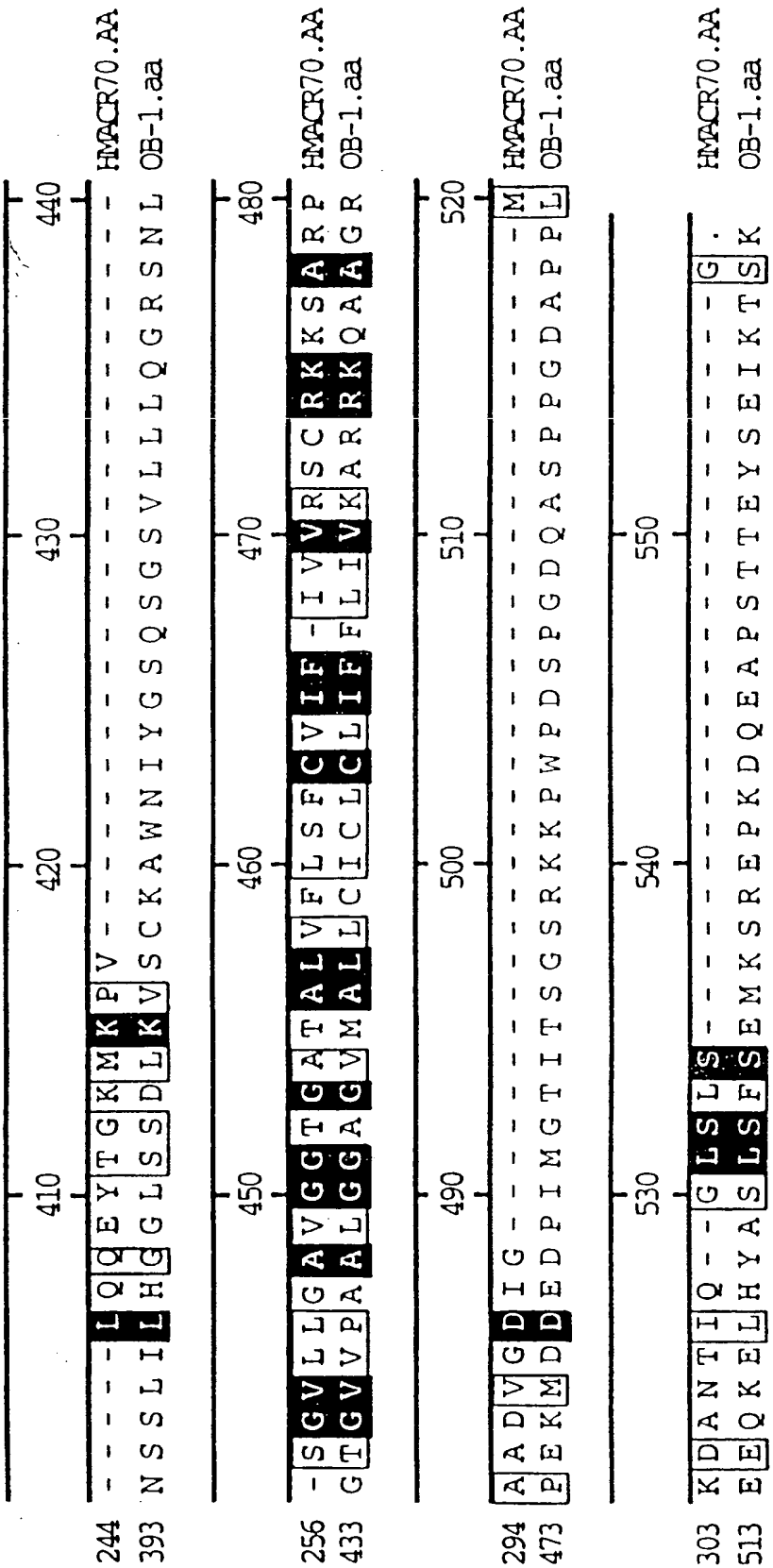


FIG.1C

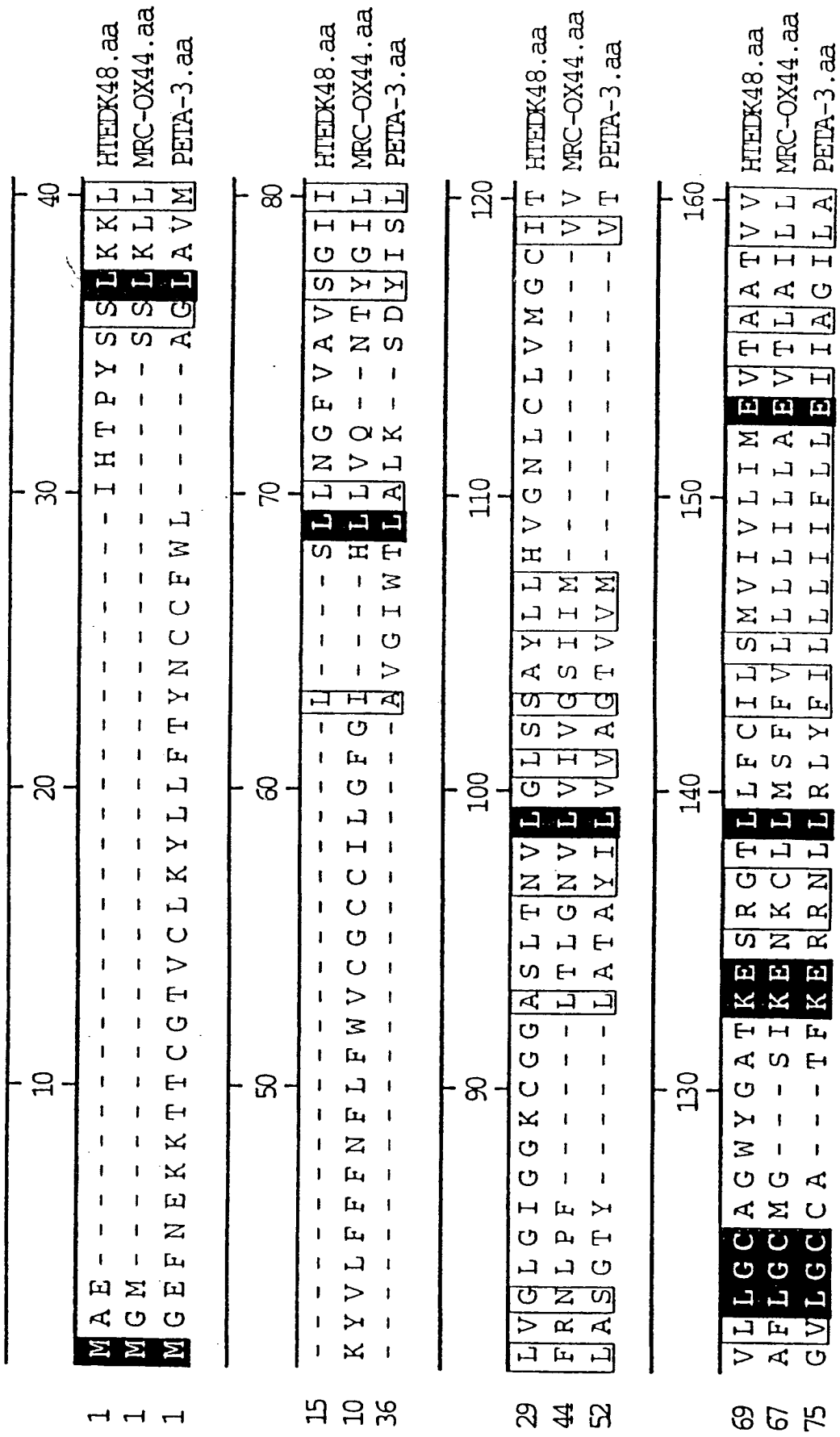


FIG.2A

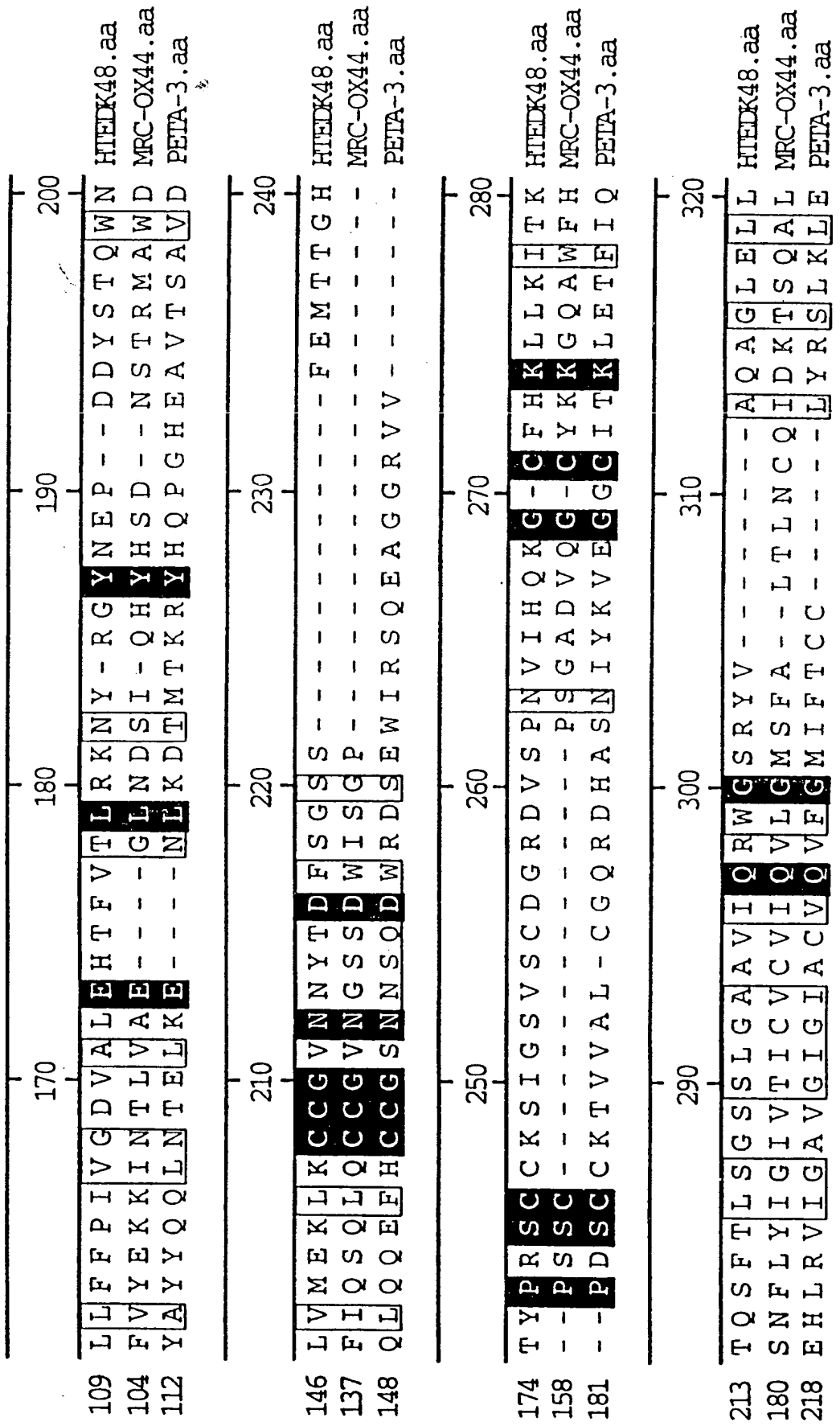


FIG.2B

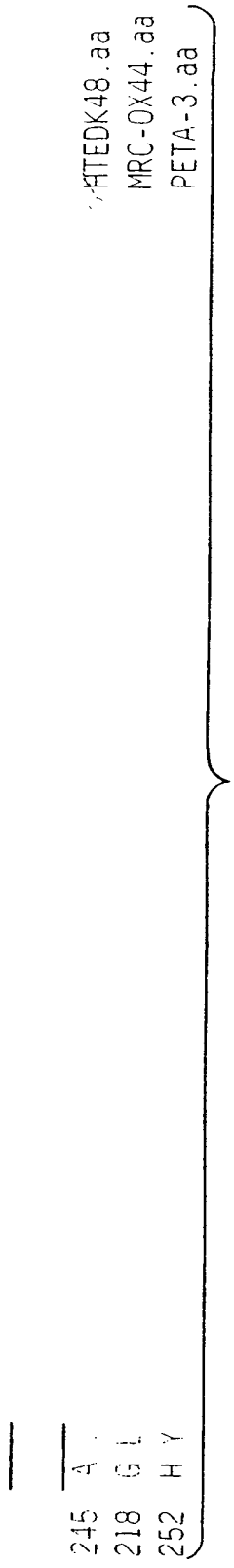


FIG.2C

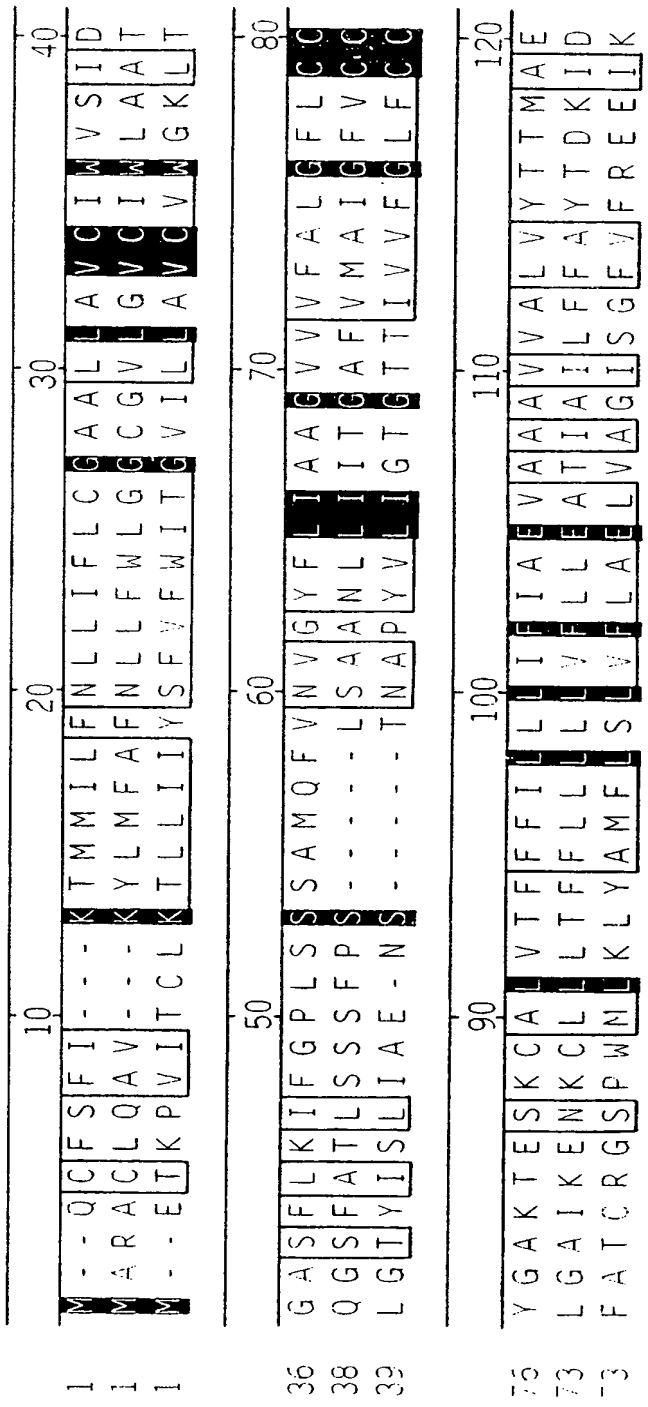


FIG.3A

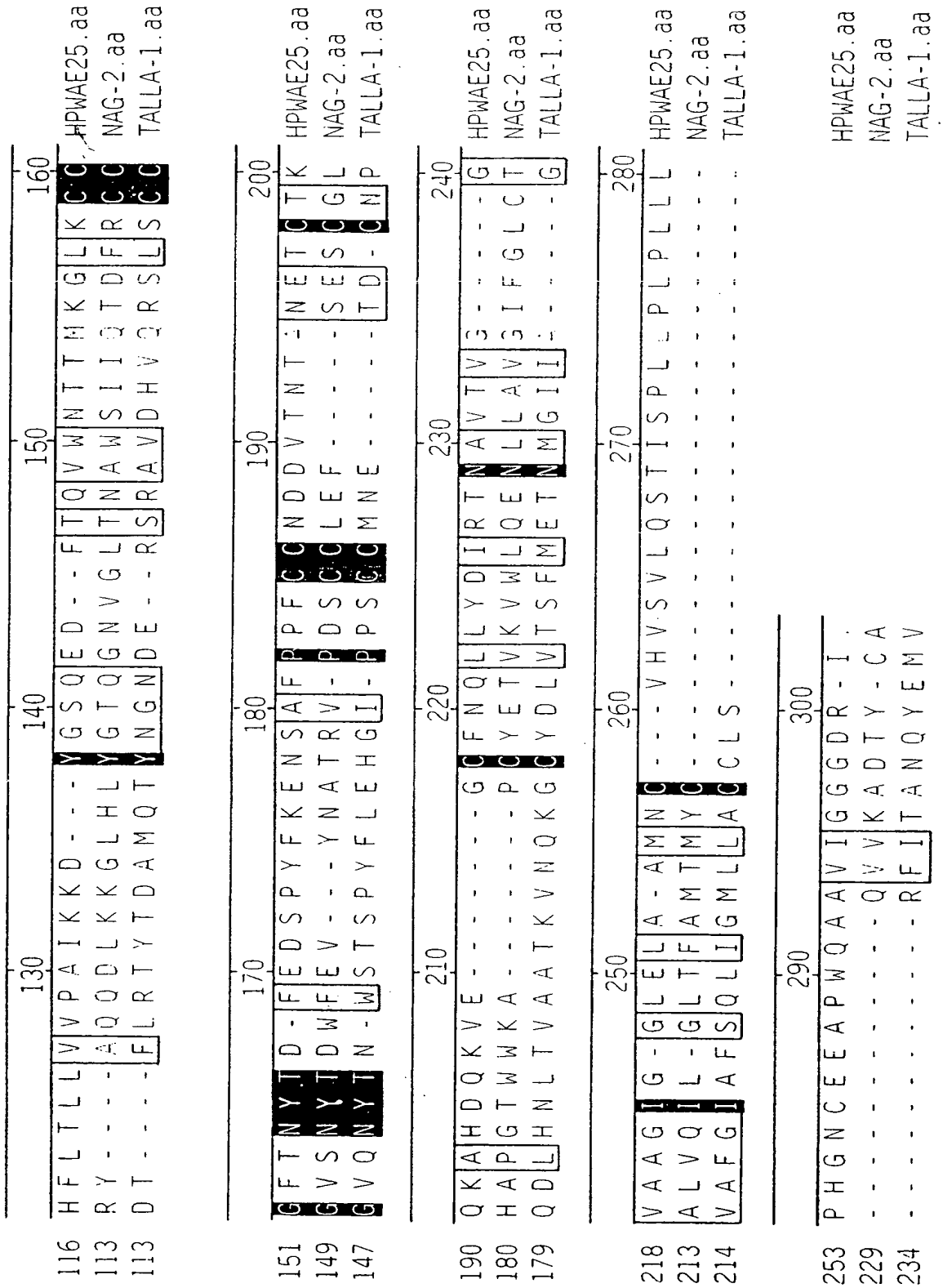


FIG.3B

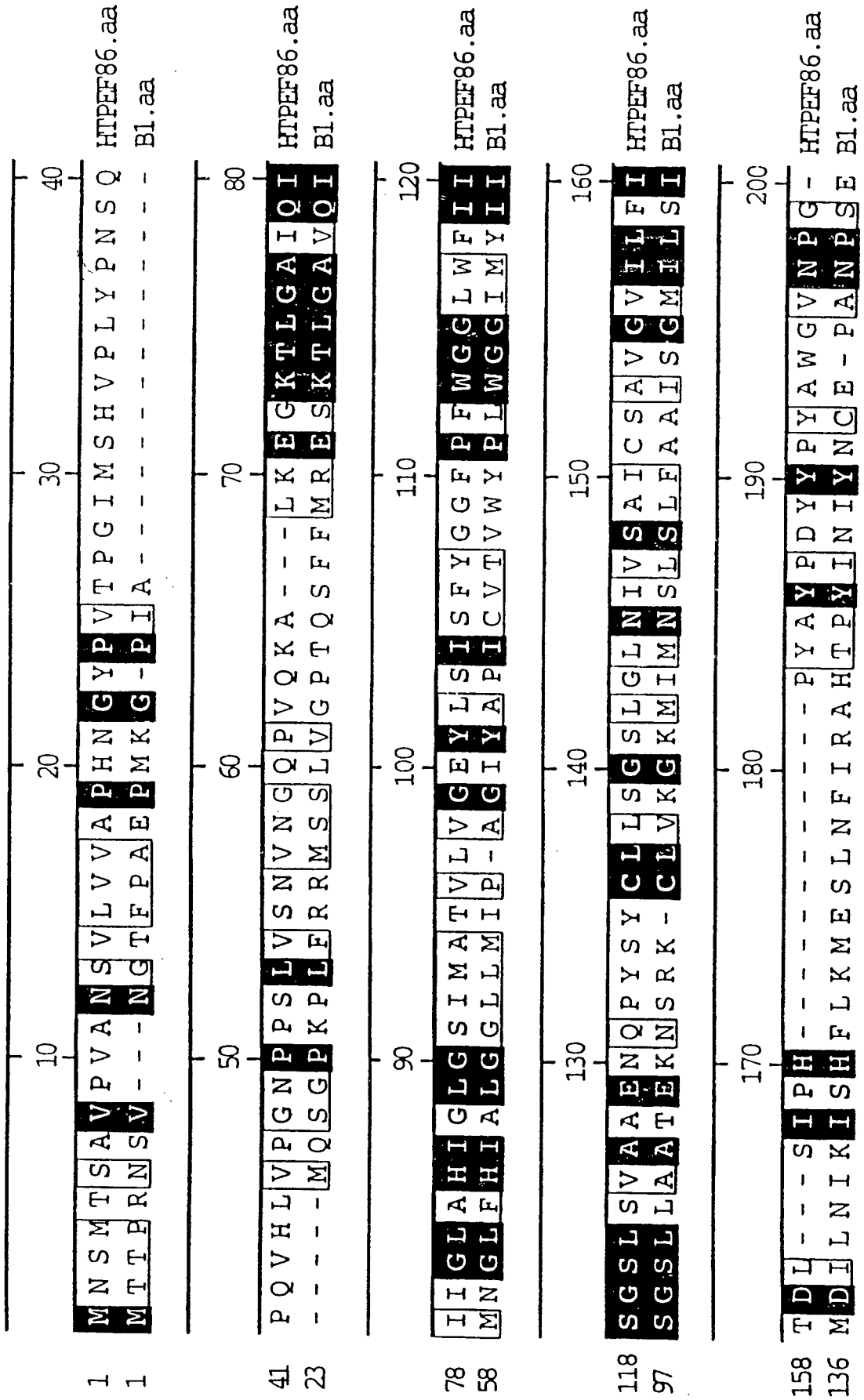


FIG.4A

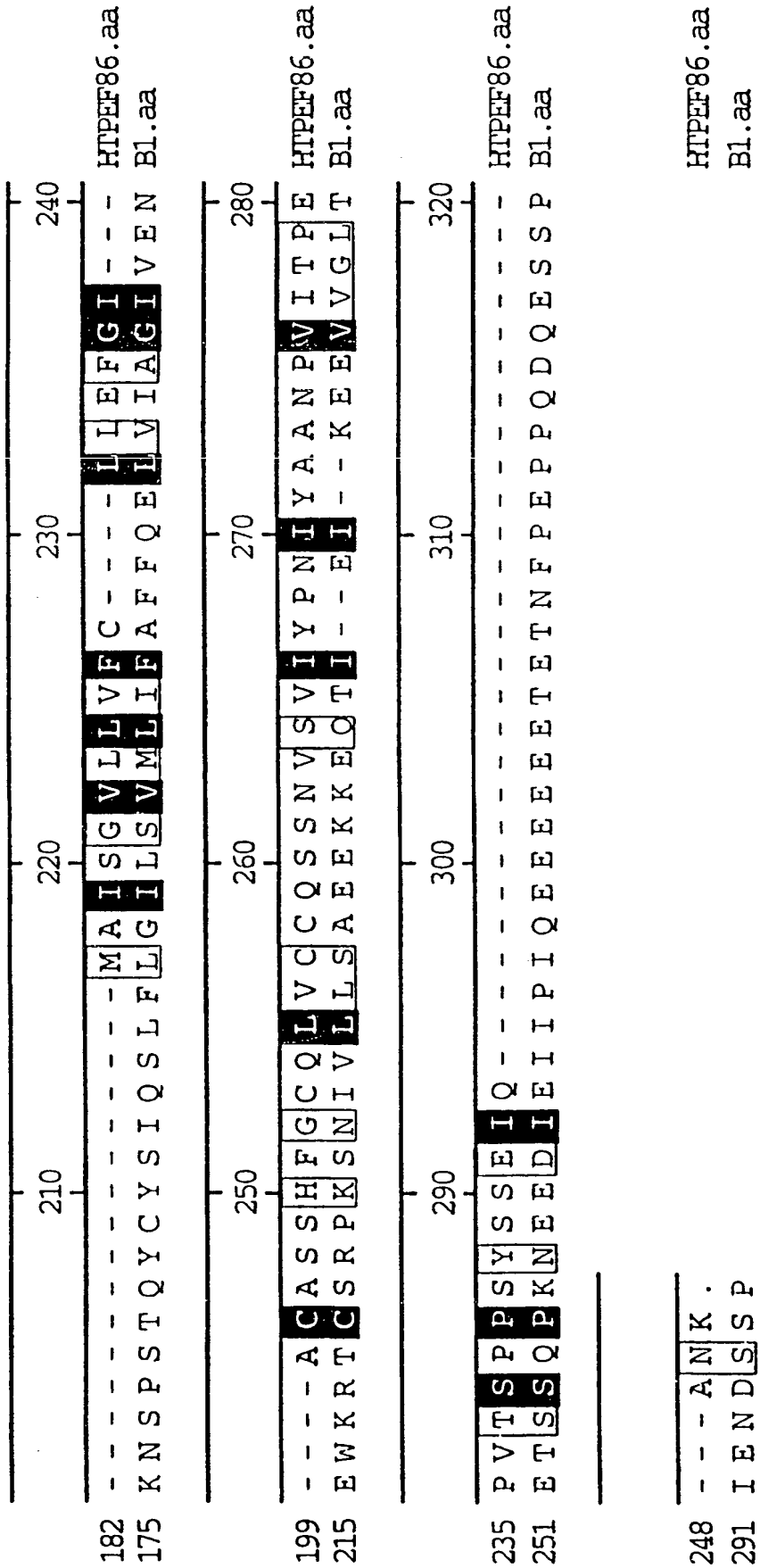


FIG.4B

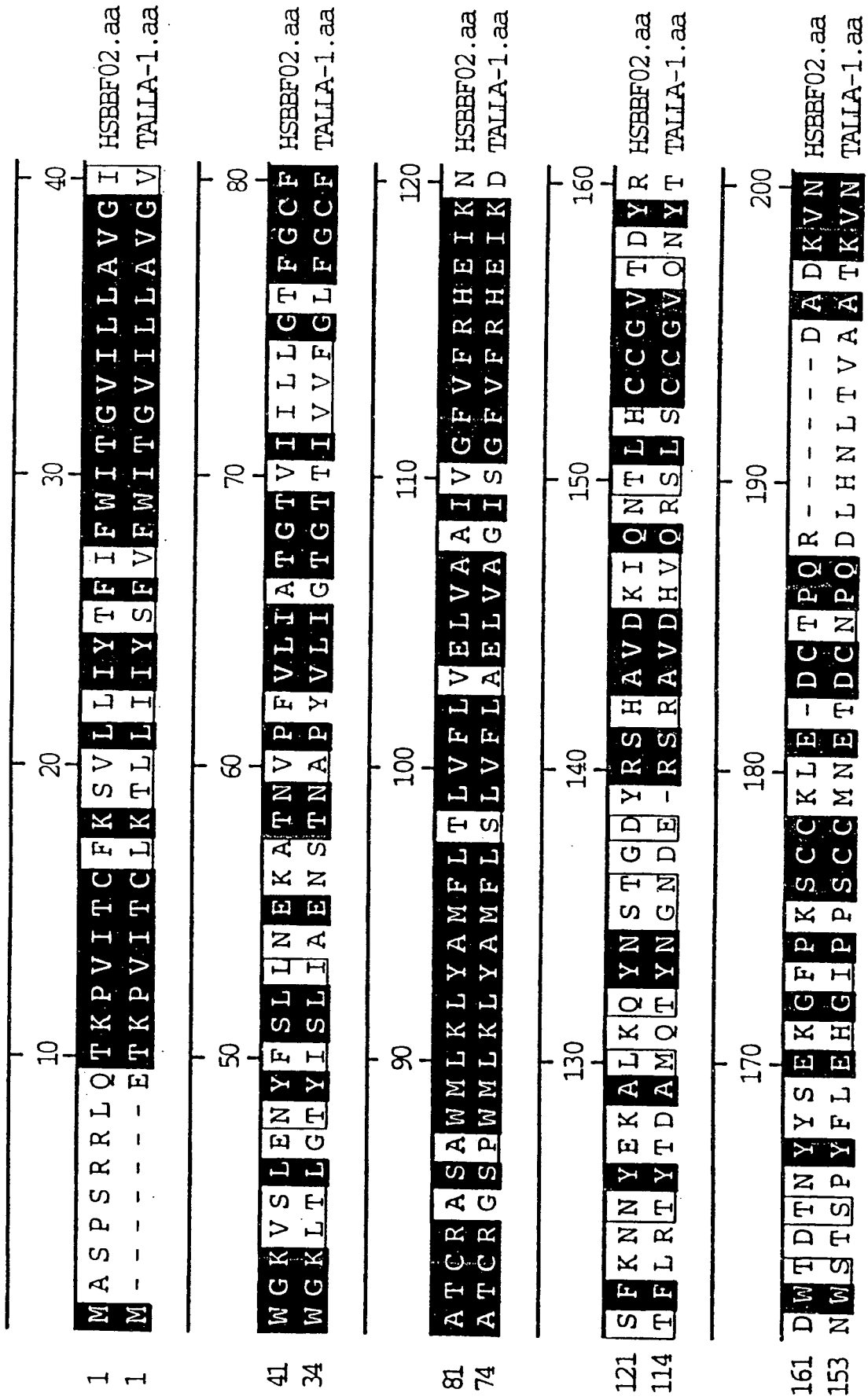


FIG.5A

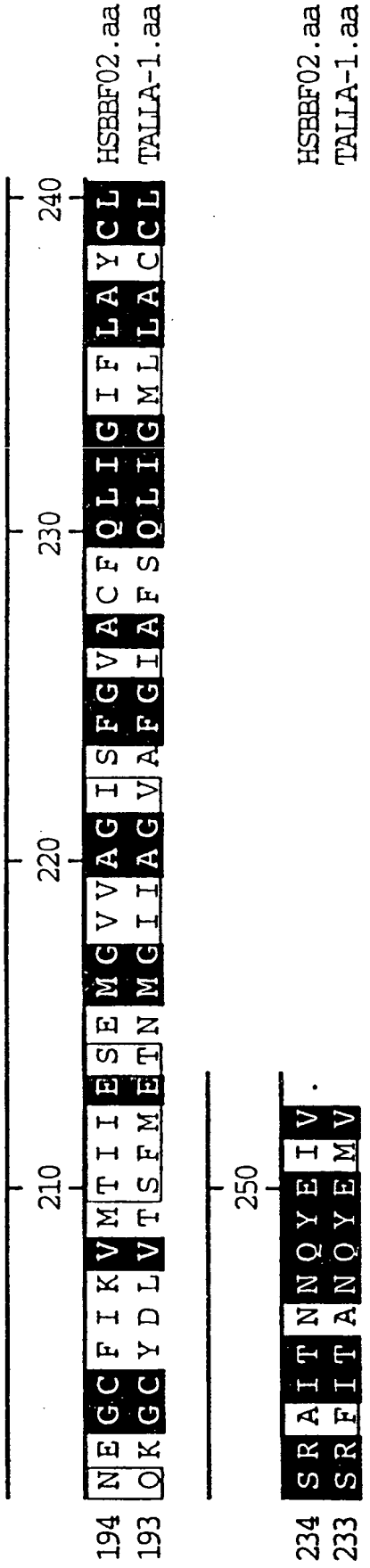


FIG.5B

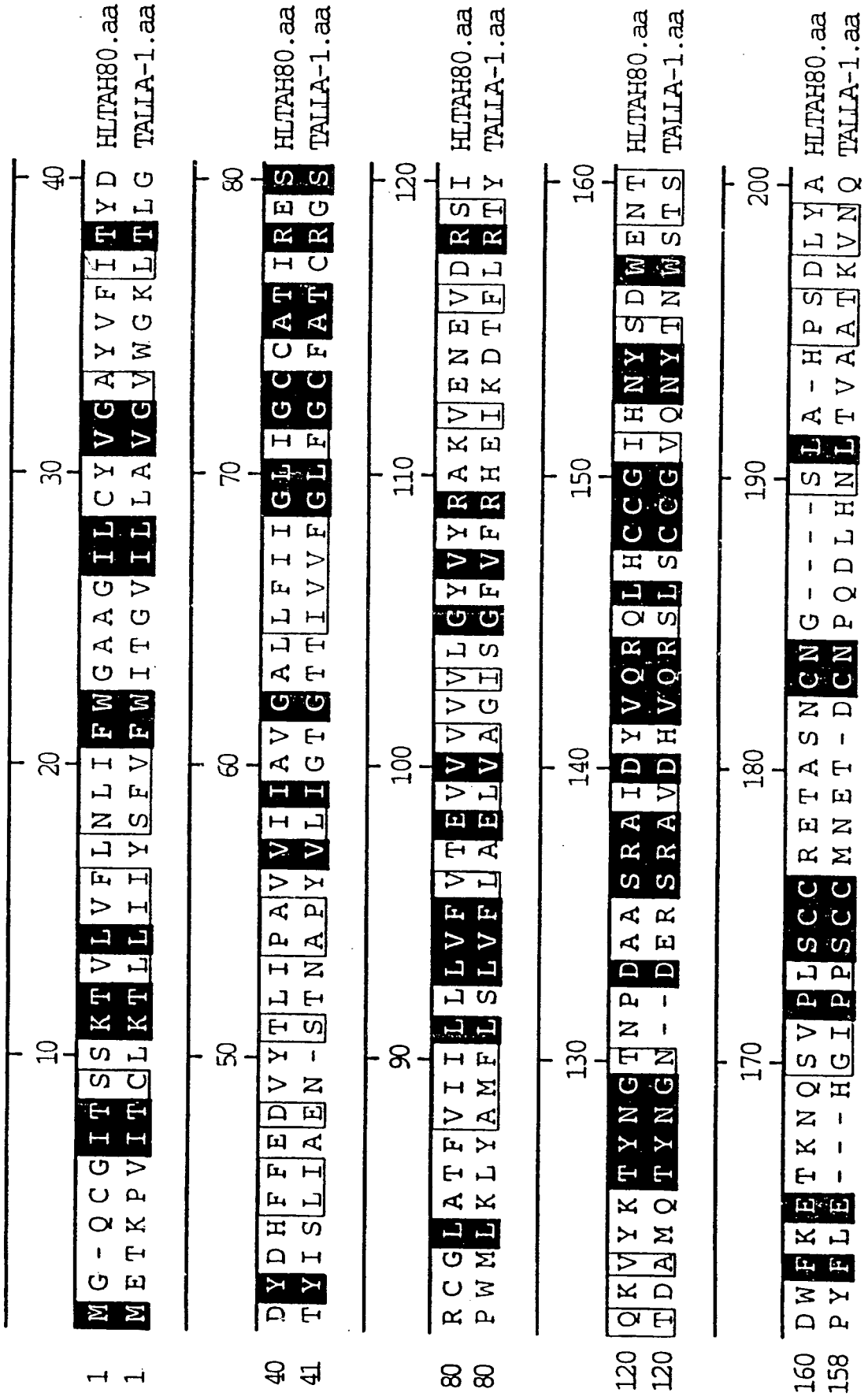


FIG.6A

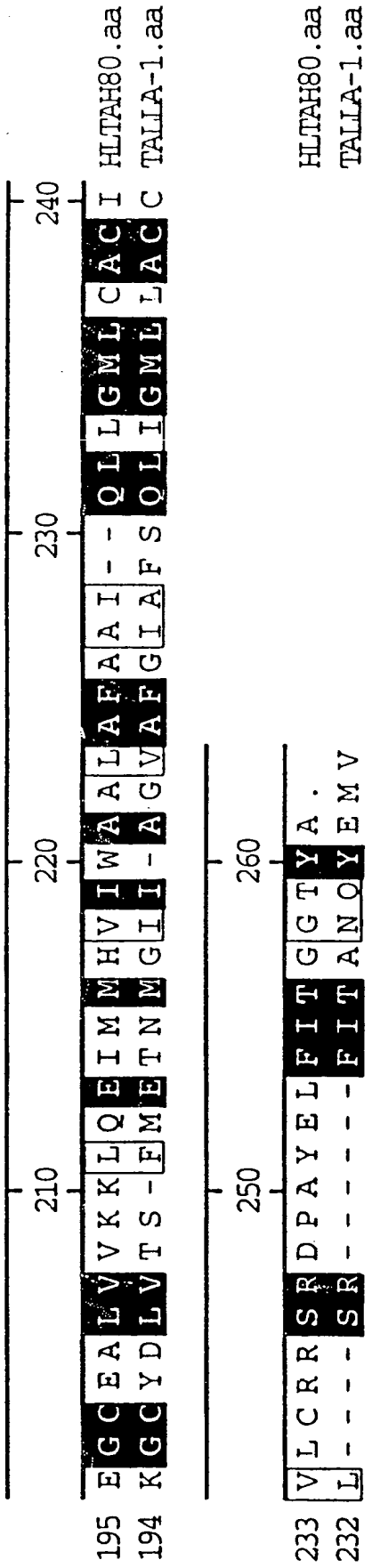


FIG.6B

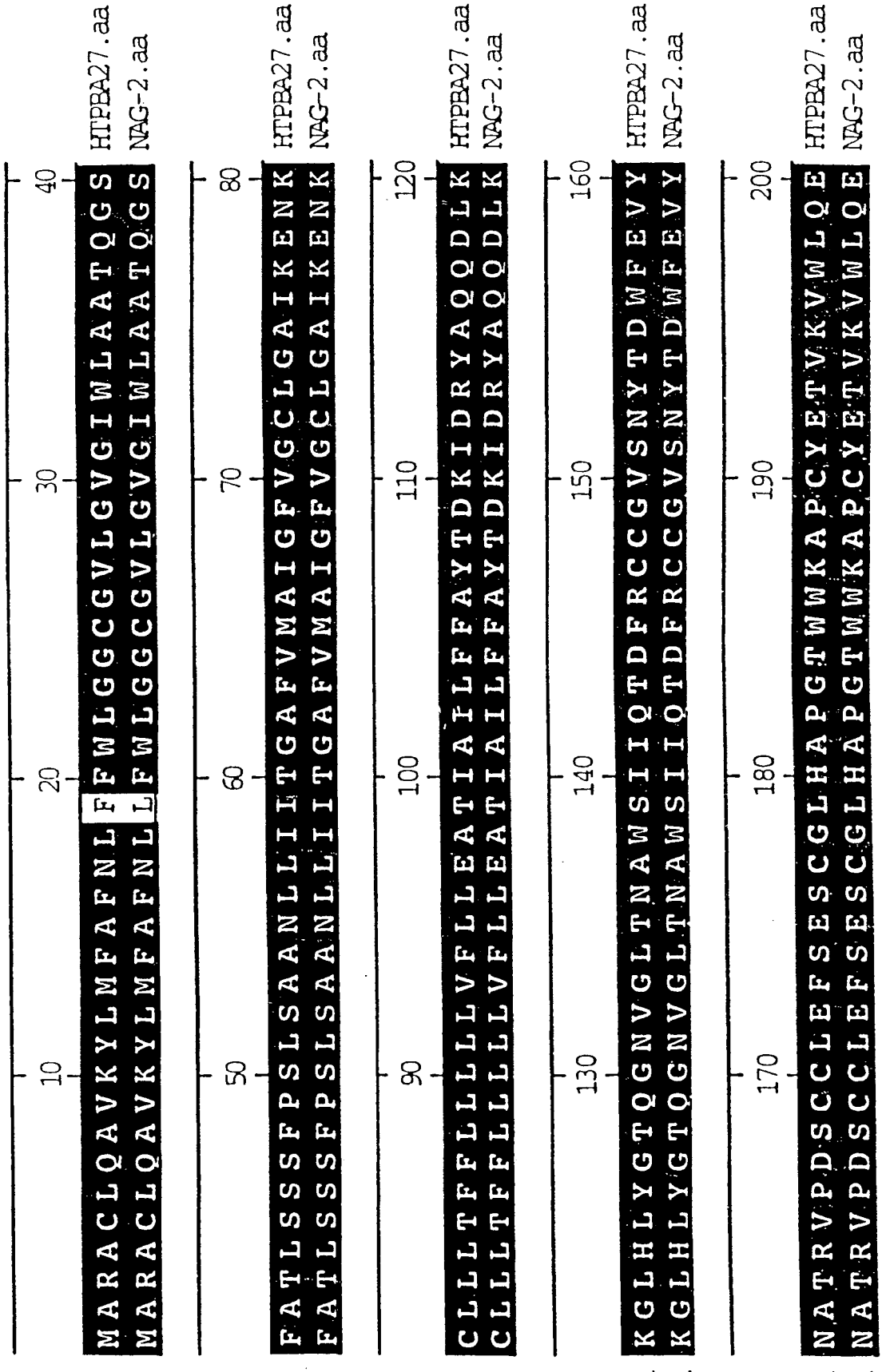

HTPBA27.aa
NAG-2.aa
MARACLQAVKYLMFAFNLFFWLGGCGVLGVGIWLAATQGS
MARACLQAVKYLMFAFNLLFWLGGCGVLGVGIWLAATQGS
FATLSSSFPSLSAANLLIITGAFVMAIGFVGCLGAIKENK
FATLSSSFPSLSAANLLIITGAFVMAIGFVGCLGAIKENK
CLLLTFFLLLLVFLLEATIAILFFAYTDKIDRYAQQDLK
CLLLTFFLLLLVFLLEATIAILFFAYTDKIDRYAQQDLK
KGLHLYGTQGNVGLTNAWSIIQTDFRCCGVSNYTDWFEVY
KGLHLYGTQGNVGLTNAWSIIQTDFRCCGVSNYTDWFEVY
NATRVPDSCCLEFSESCGLHAPGTWWKAPCYETVKVWLQE
NATRVPDSCCLEFSESCGLHAPGTWWKAPCYETVKVWLQE

FIG.7A

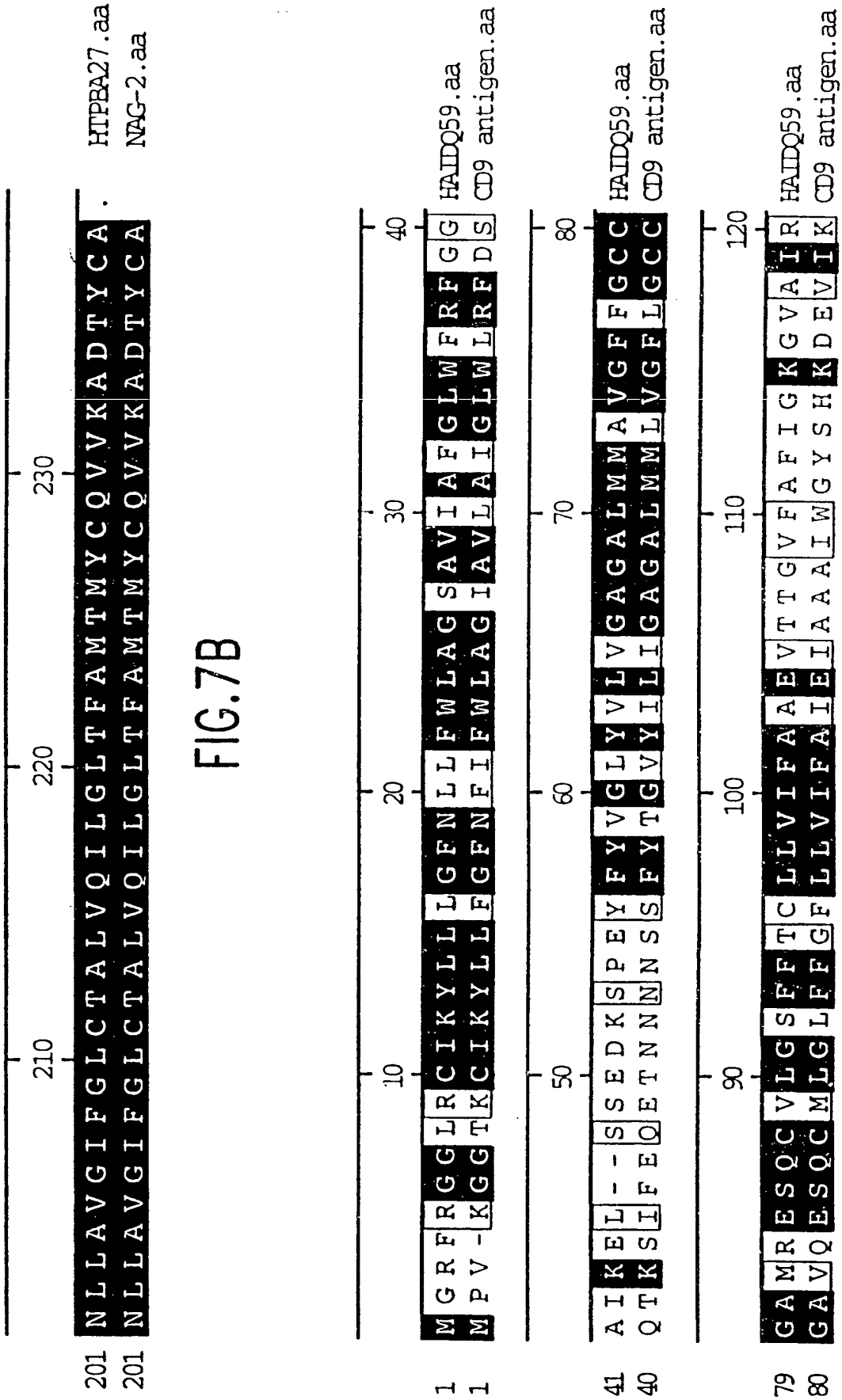


FIG.7B

FIG.8A

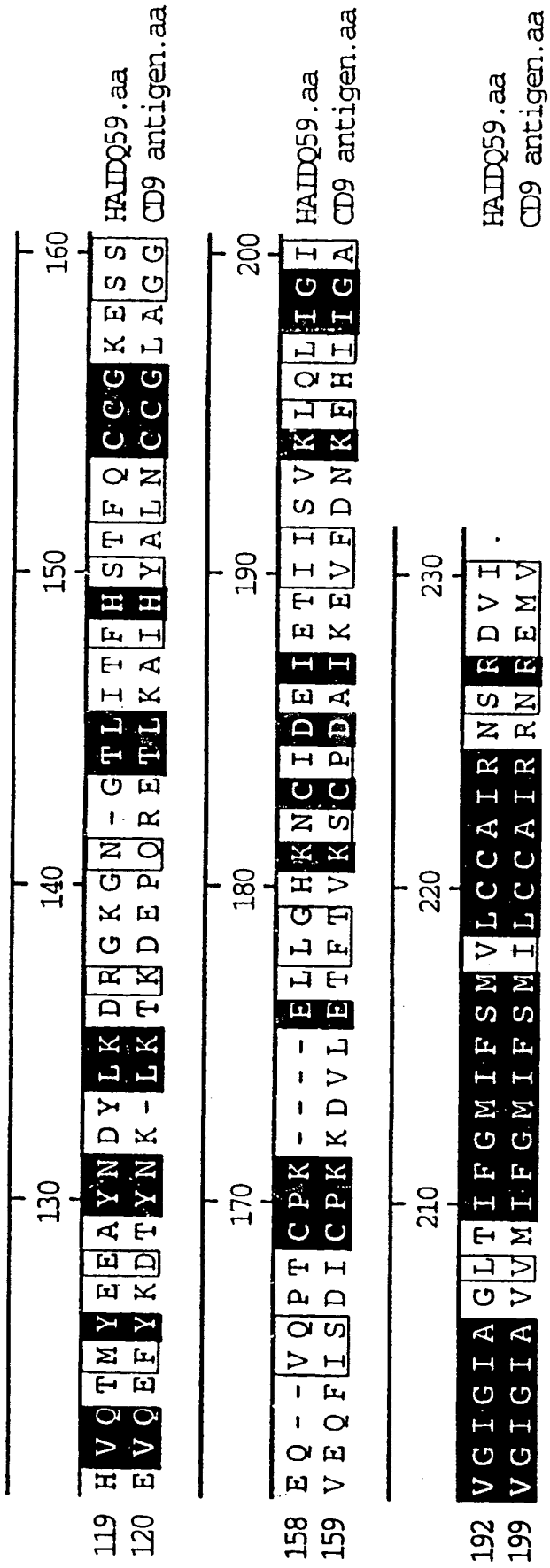


FIG.8B

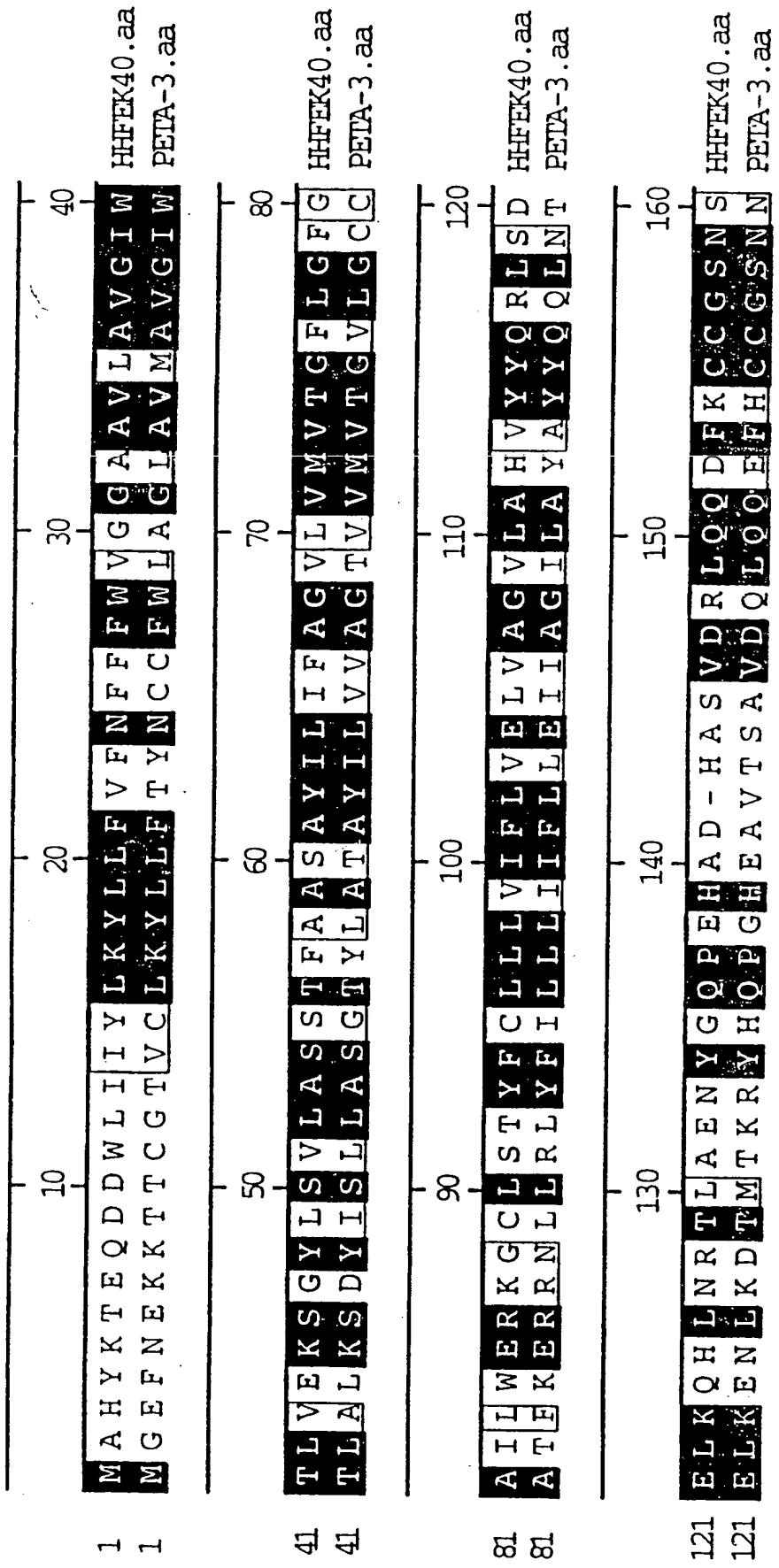


FIG.9A

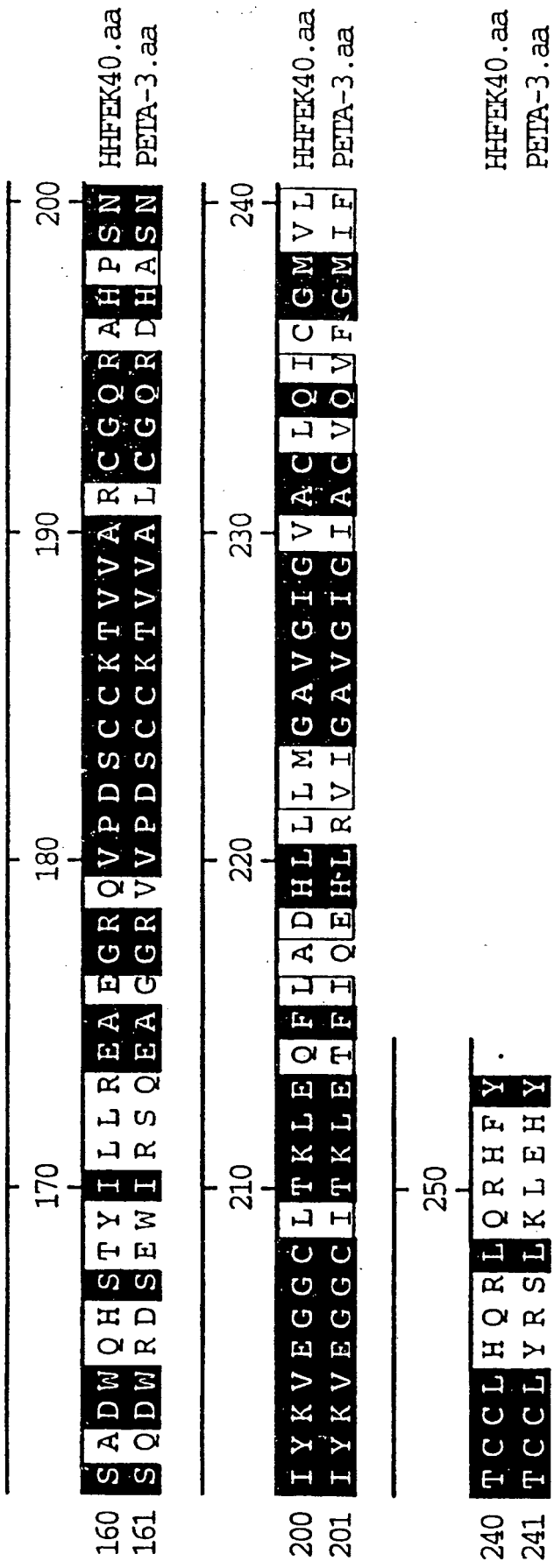


FIG.9B

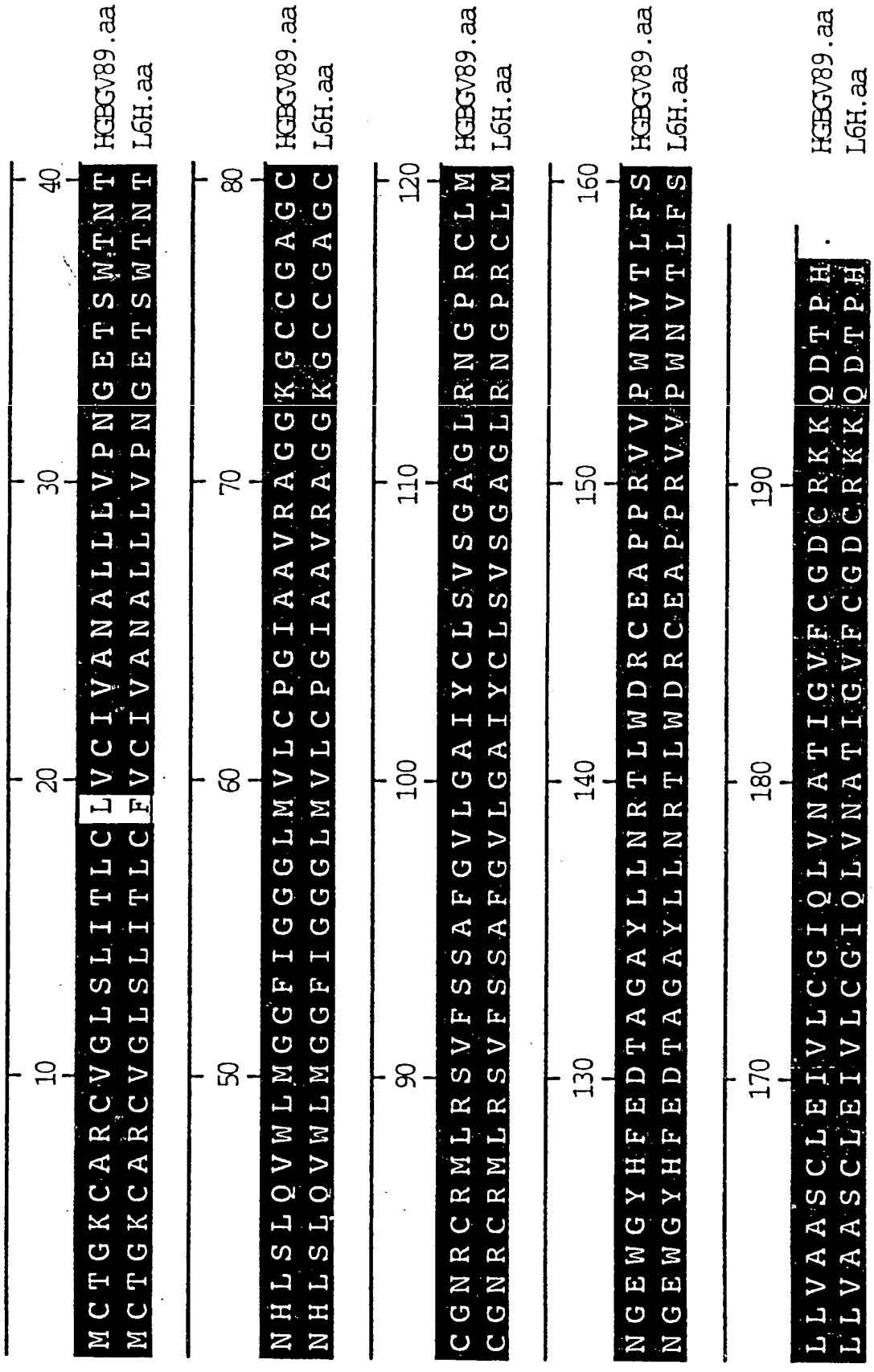


FIG.10

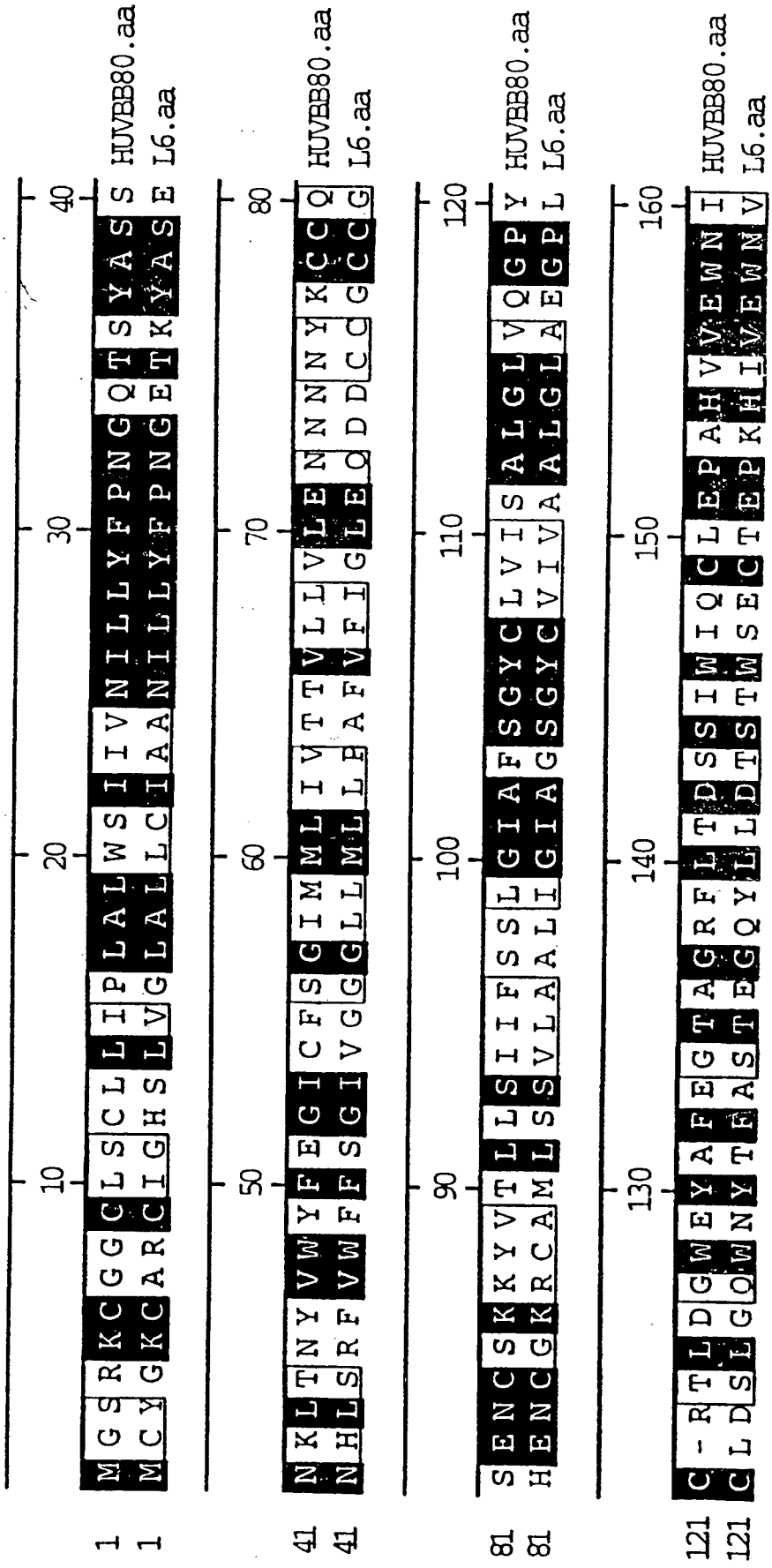


FIG.11A

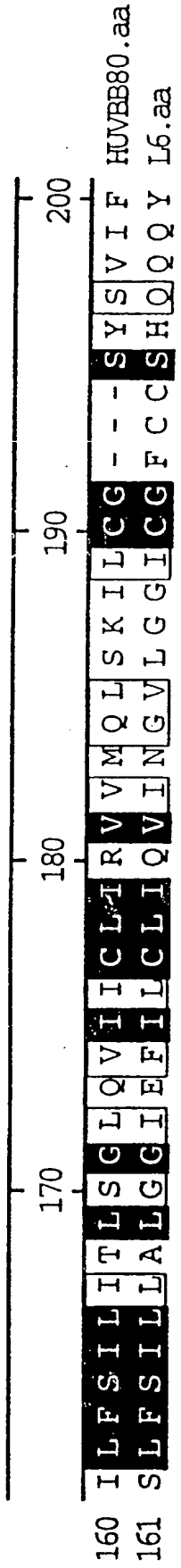


FIG.11B

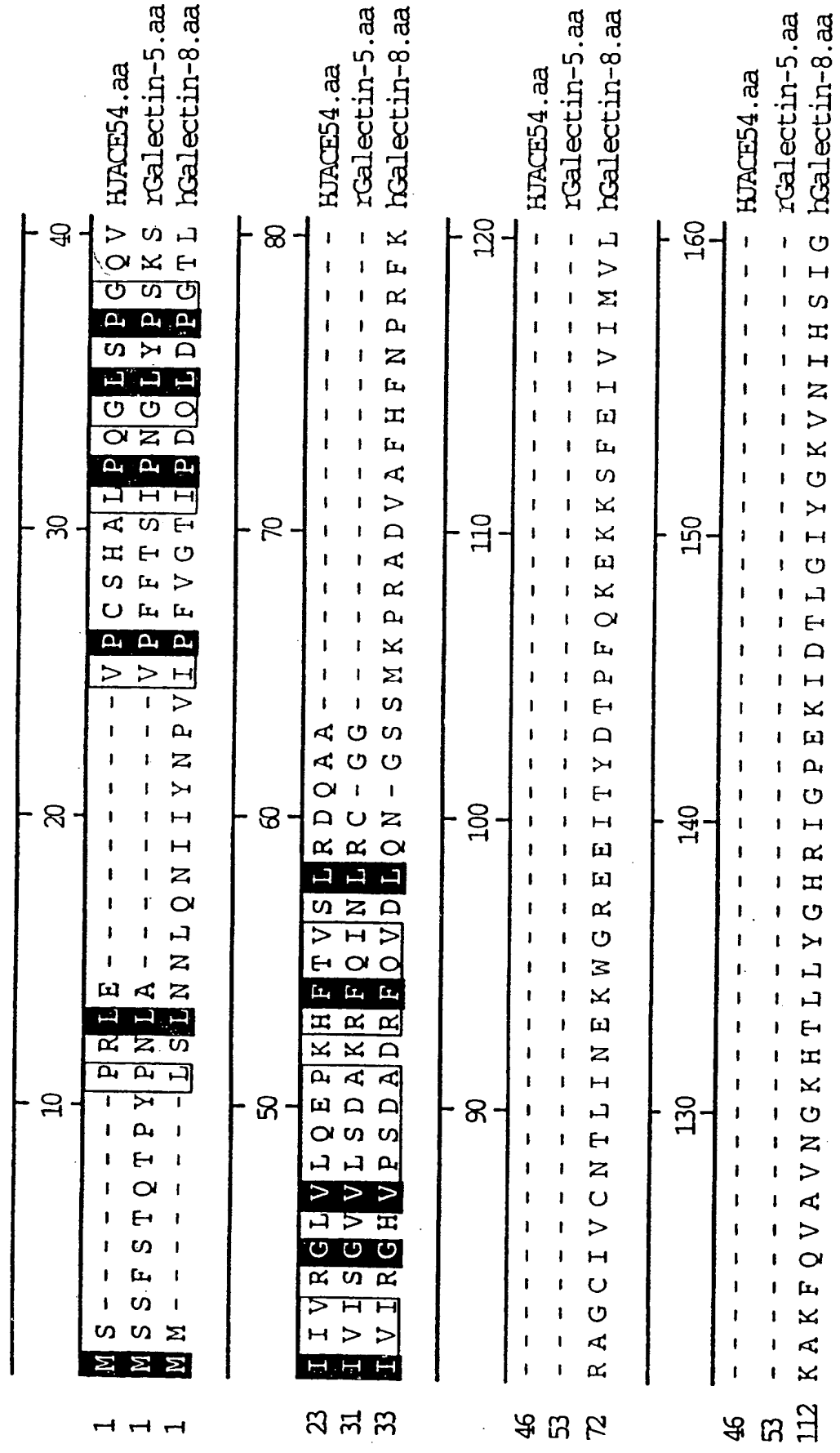


FIG.12A

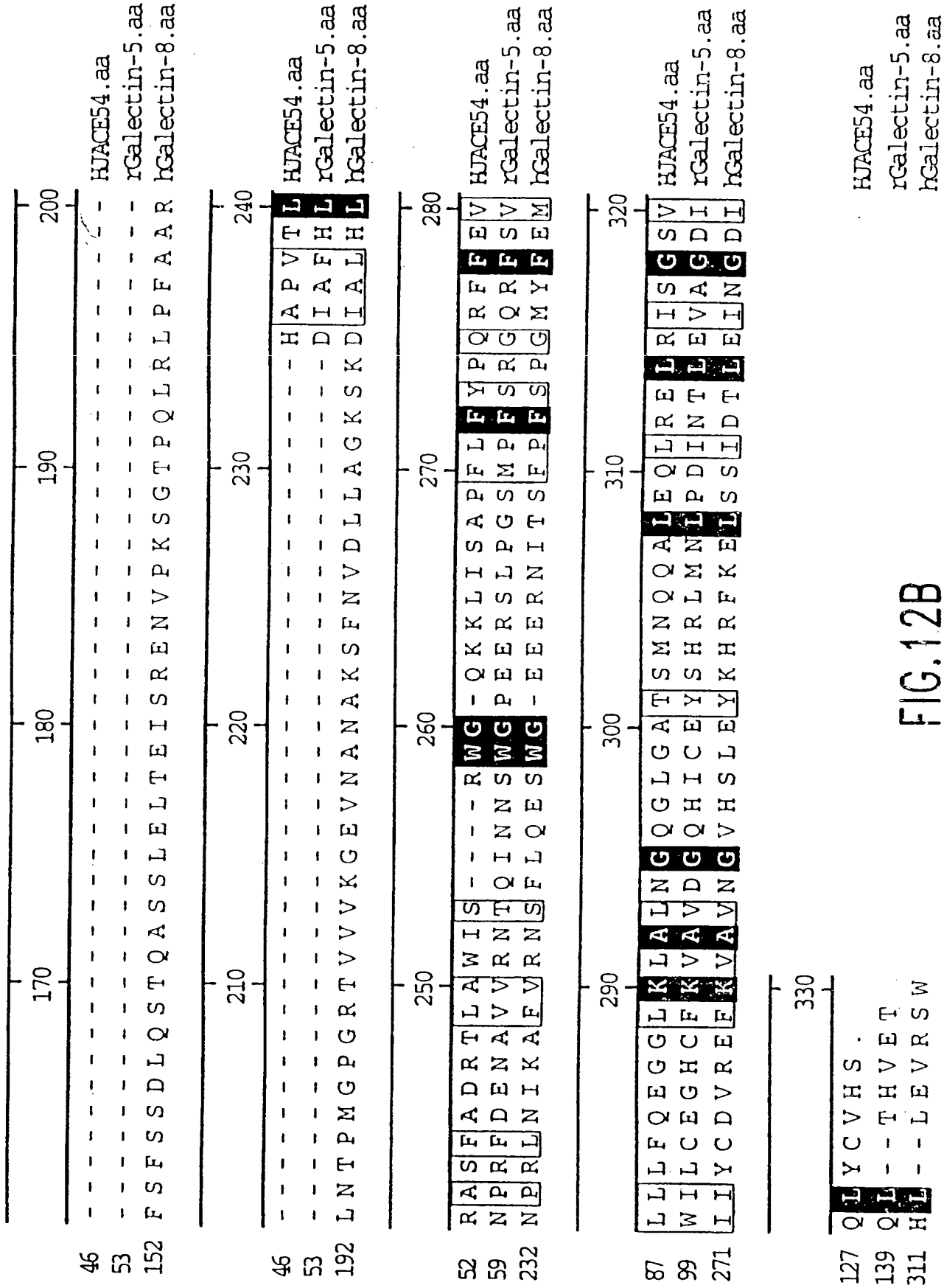


FIG.12B

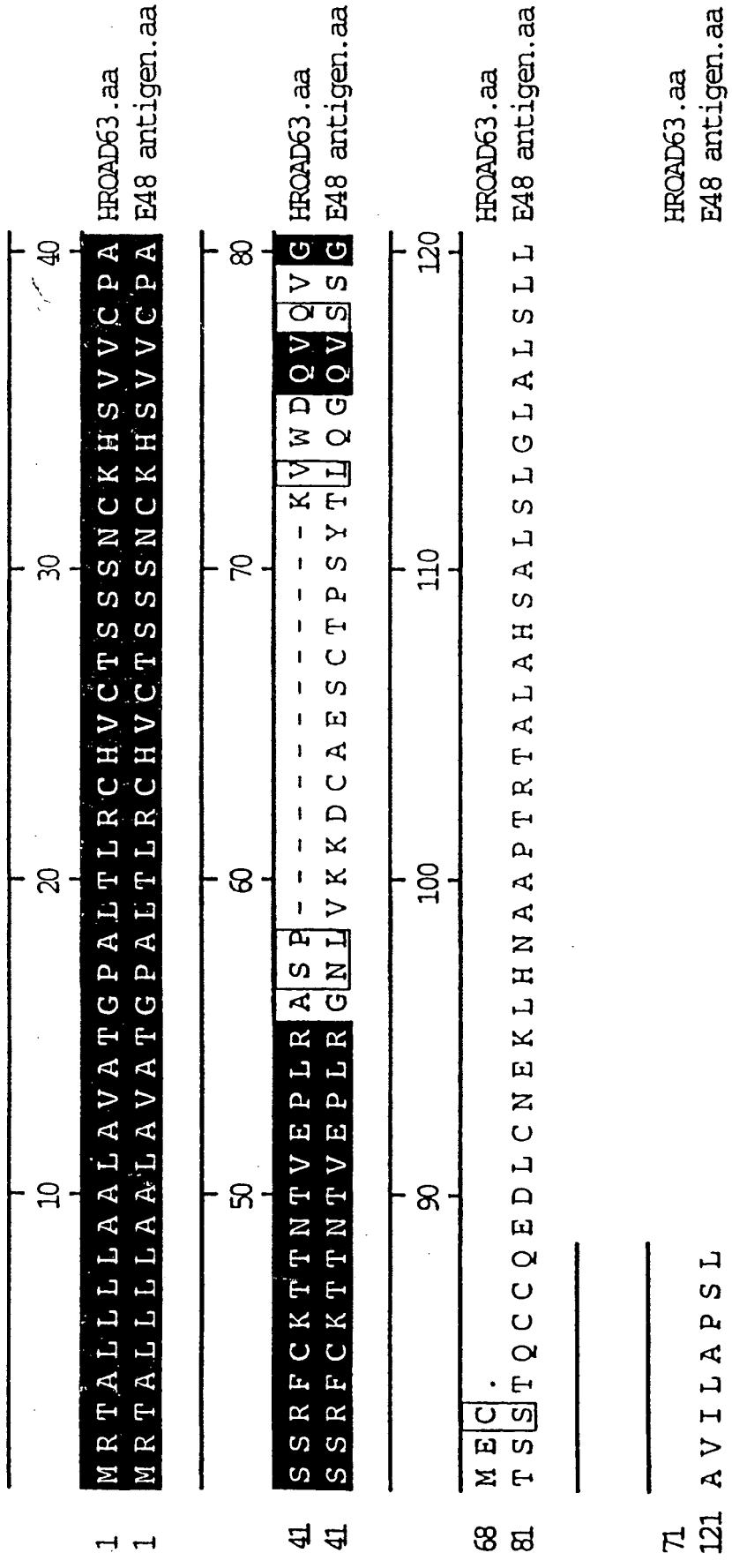


FIG.13

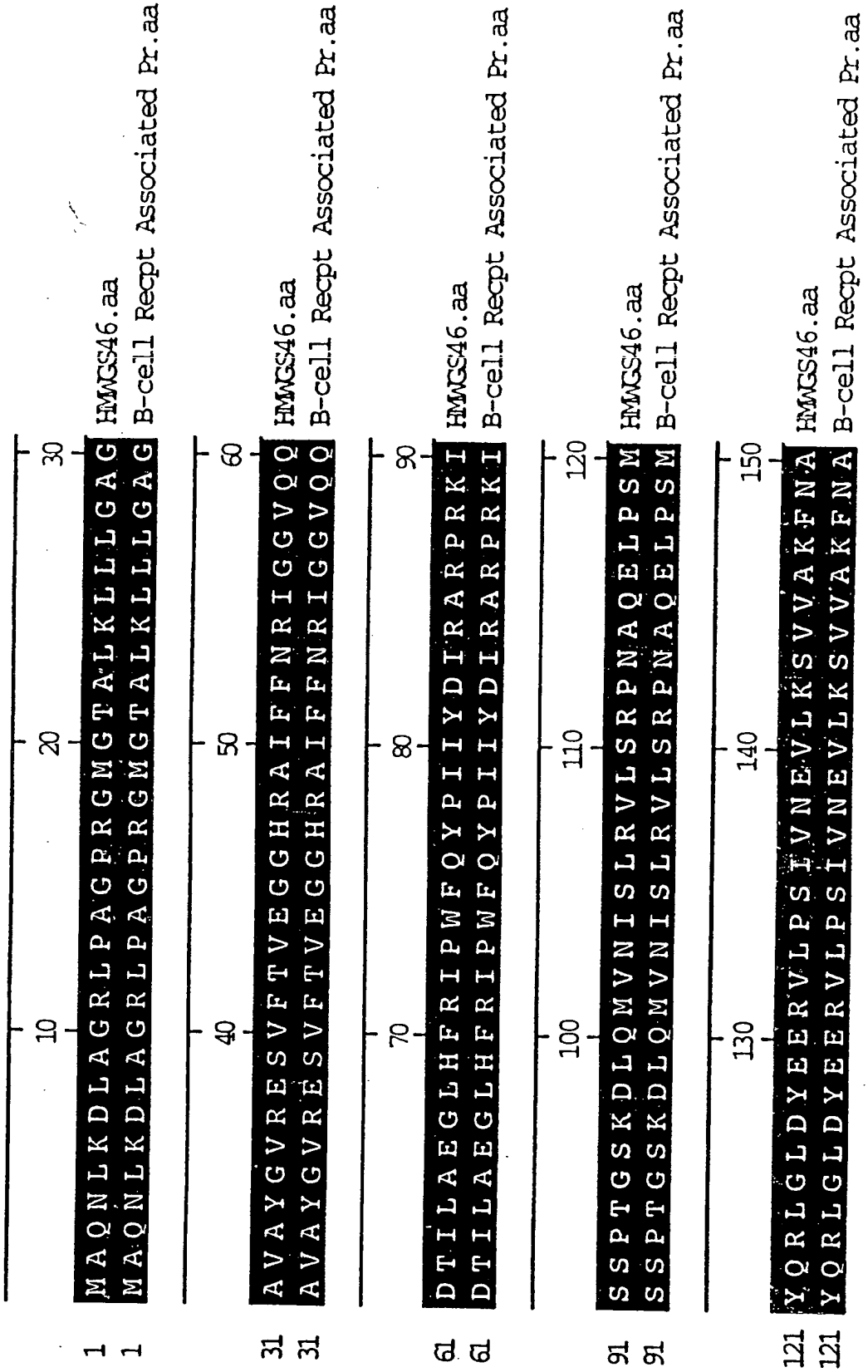


FIG.14A

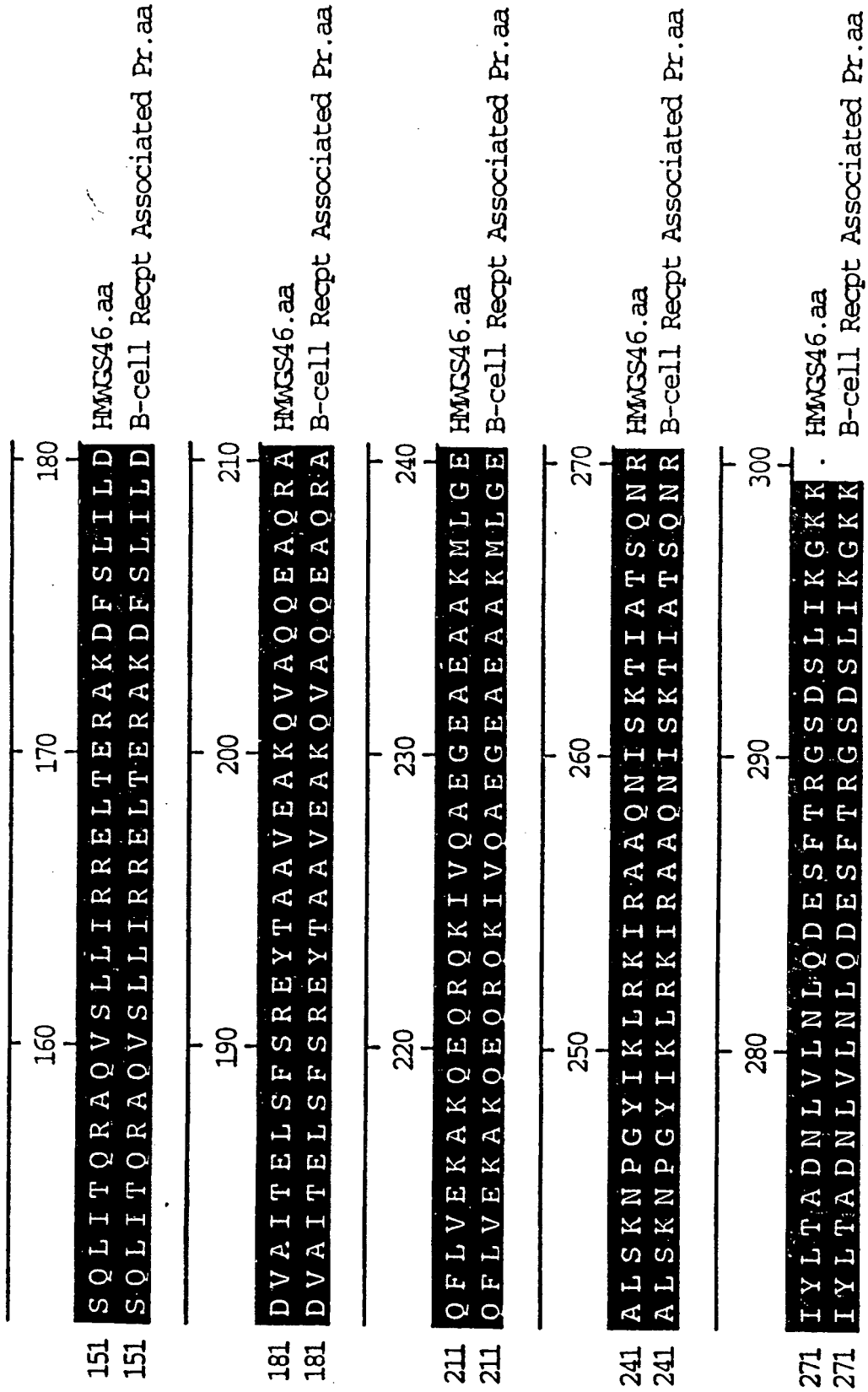


FIG.14B

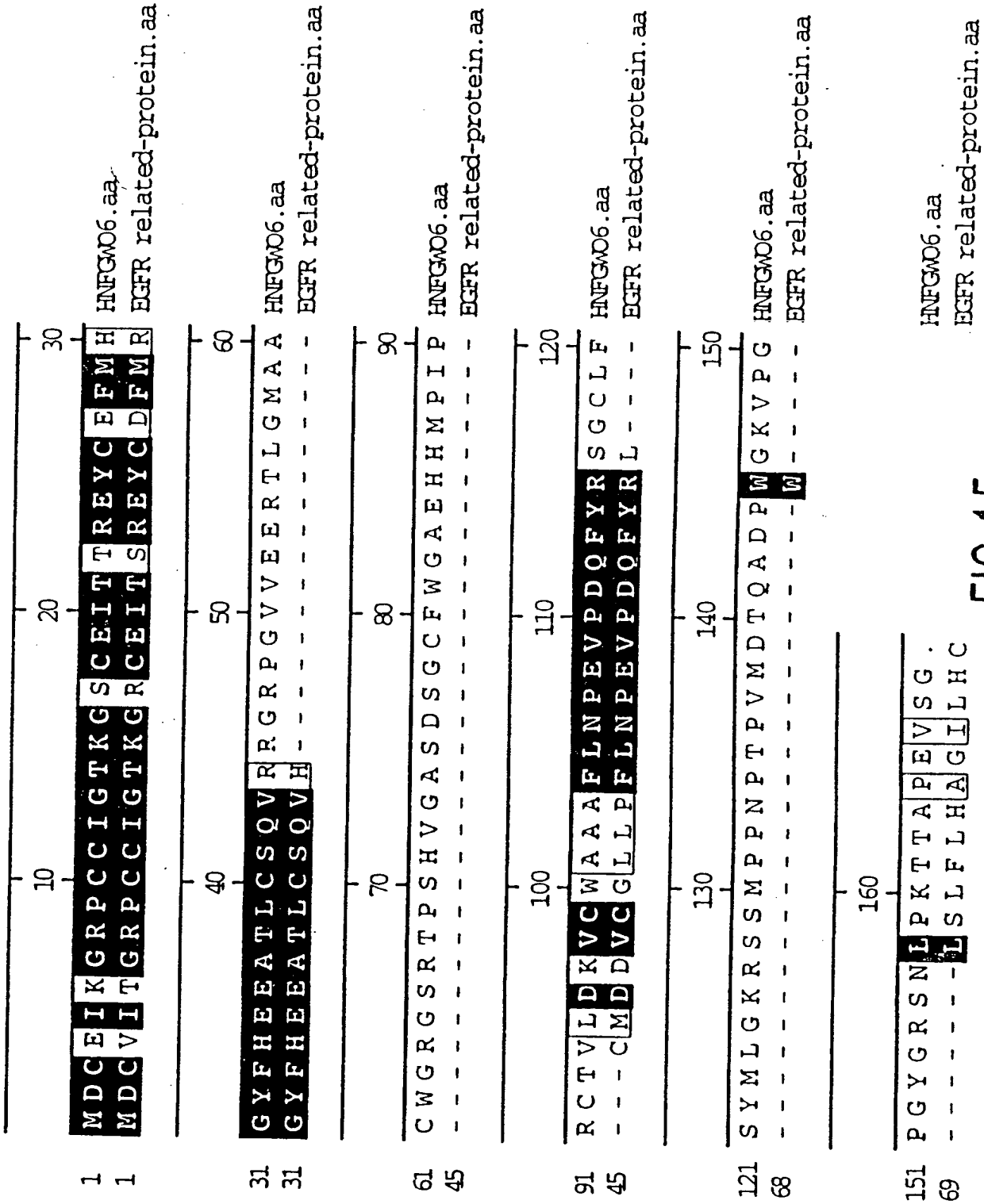


FIG.15

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/00959

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C12N15/85 C12N5/10 C07K14/705 C07K16/28
C12Q1/68 A61K38/17 A61K48/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K C12Q A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE GENBANK Accession No. T48852, 8 February 1995 "EST70319 from Stratagene placenta" XP002066898 see the whole document ---	1-12,20
X	DATABASE GENBANK Accession No. G14442, 4 January 1996 "STS SHGC-10055" XP000206689 see the whole document ---	1-3,20
	-/--	

[X] Further documents are listed in the continuation of box C.

[ ] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
4 June 1998	16. 09. 1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Galli, I

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ZANNETTINO A.C.W. ET AL.: "A powerful technique for isolating genes encoding surface antigens using retroviral expression cloning" J. IMMUNOL., vol. 156, no. 2, 15 January 1996, pages 611-620, XP002066893 see abstract	1-21
A	ADAMS M D ET AL: "COMPEMENTARY DNA SEQUENCING: EXPRESSED SEQUENCE TAGS AND HUMAN GENOME PROJECT" SCIENCE, vol. 252, no. 5013, 21 June 1991, pages 1651-1656, XP000645049 cited in the application see abstract	1-21
A	SIMMONS D. & SEED B.: "Isolation of a cDNA encoding CD33, a differentiation antigen of myeloid progenitor cells" J. IMMUNOLOGY, vol. 141, no. 8, 15 October 1988, pages 2797-2800, XP002066894 see abstract see figure 4	1-21
A	CROCKER P.R. ET AL.: "Sialoadhesin, a macrophage sialic acid binding receptor for haemopoietic cells with 17 immunoglobulin-like domains" EMBO J., vol. 13, no. 19, 3 October 1994, pages 4490-4503, XP002066895 see abstract see figure 2	1-21
A	KELM S. ET AL.: "The Sialoadhesins: a family of sialic acid-dependent cellular recognition molecules within the immunoglobulin superfamily" GLYCOCONJ. J., vol. 13, no. 6, December 1996, pages 913-926, XP002066896 see the whole document	1-21
A	DATABASE GENBANK Accession No. U71382, 12 November 1996 PATEL N. ET AL.: "OB binding protein - 1" XP002066901 cited in the application see the whole document	1-21

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	DATABASE GENBANK Accession No. AA344713, 21 April 1997 "EST50650, H. sapiens cDNA similar to CD33 antigen" XP002066900 see the whole document ---	1-12,20
P,A	TAKEI Y. ET AL.: "Molecular cloning of a novel gene similar to myeloid antigen CD33 and its specific expression in placenta" CYTOGENET. CELL GENET., vol. 78, 1997, pages 295-300, XP002066897 see abstract see figure 1 -----	1-21

INTERNATIONAL SEARCH REPORT

Inter onal application No.
PCT/US 98/00959

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Remark: Although claims 13 and 14 are directed to a method of treatment of the human/animal body, and claim 15 (as far as in vivo methods are envisaged) is directed to a diagnostic method, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see further information sheet

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [X] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

see further information sheet, subject 1.

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International Application No. PCT/ US 98/00959

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: (1-21) - partial

An isolated polynucleotide encoding the HMACR70 cell surface receptor (Seq. IDs 1,18), polynucleotides complementary thereto, fragments thereof, homologs thereof.

An expression system comprising said polynucleotides, a process to produce a cell harbouring said expression system. Said cell or membrane thereof expressing said receptor.

A process for the production of the HMACR70 polypeptide comprising culturing said cell.

Said HMACR70 polypeptide or homologs. An antibody directed against said polypeptide.

Therapeutic and diagnostic applications.

A method to identify agonists or antagonists to said receptors. Said agonists and antagonists.

2. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 2,3,19

3. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 4,20

4. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 5,21

5. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 6,22

6. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 7,23

7. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 8,24

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

8. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 9,10,25

9. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 11,26

10. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 12,27

11. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 13,28

12. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 14,29

13. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 15,30

14. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 16,31

15. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 17,21

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, 15/85, 5/10, C07K 14/705, 16/28, C12Q 1/68, A61K 31/70, 38/17, 48/00

A2

(11) International Publication Number: WO 98/31799

(43) International Publication Date: 23 July 1998 (23.07.98)

(21) International Application Number: PCT/US98/00959

(22) International Filing Date: 21 January 1998 (21.01.98)

(30) Priority Data:
60/034,204 21 January 1997 (21.01.97) US
60/034,205 21 January 1997 (21.01.97) US

(71) Applicant *(for all designated States except US)*: HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: NI, Jian [CN/US]; 5502 Manorfield Road, Rockville, MD 20853 (US). GENTZ, Reiner, L. [DE/US]; 13404 Fairland Park Drive, Silver Spring, MD 20904 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US).

(74) Agents: BROOKES, A., Anders et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: POLYNUCLEOTIDES AND POLYPEPTIDES ENCODING RECEPTORS

(57) Abstract

Receptor polypeptides and polynucleotides and methods for producing such polypeptides by recombinant techniques are disclosed. Also disclosed are methods for utilizing receptor polypeptides and polynucleotides in the design of protocols for the treatment of diseases and diagnostic assays for such conditions.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

POLYNUCLEOTIDES AND POLYPEPTIDES ENCODING RECEPTORS

FIELD OF INVENTION

This invention relates to newly identified polynucleotides and the polypeptides encoded by them, the use of such polynucleotides and polypeptides, and their production. More particularly, the polynucleotides and polypeptides of the present invention relate to specific receptor families described in the specification and known in the art. The invention also relates to inhibiting or activating the action of such polynucleotides and polypeptides.

BACKGROUND OF THE INVENTION

Receptor proteins are found on the membrane of the cells and are generally involved in signal transduction. There are many types of receptor proteins, and for convenience, these proteins are grouped in families based on similarity in structure and function.

For example, the TM4SF superfamily of cell surface proteins, also known as the tetraspan receptor superfamily, is comprised of at least seventeen individual gene products (these include CD9, CD20, CD37, CD53, CD63, CD81, CD82, A15, CO-029, Sm23, RDS, Uro B, Uro A, SAS, Rom-1, PETA3, and YKK8). The TM4SF superfamily is the second largest group in the CD antigen superfamily. Each member of the TM4SF superfamily can be characterized by several putative physical features including four highly conserved transmembrane domains, two divergent extracellular loops, and two short and highly divergent cytoplasmic tails. Expression patterns for members of the TM4SF superfamily tend to be rather broad and can vary widely between members. The functional roles of TM4SF superfamily members are primarily associated with signal transduction events and pathways, but also include cell adhesion in platelets and other lymphocytic and non-lymphocytic cell lines, as well as cell motility, proliferation, and metastasis. In addition, recent evidence suggests that a subset of the members of the TM4SF superfamily may function as potassium channel molecules.

One member of the TM4SF family, CD20, is a four membrane spanning domain cell surface phosphoprotein expressed exclusively on B lymphocytes. Although the precise functional role of CD20 has yet to be determined, it is thought to function primarily as a receptor during B-cell activation. Furthermore, a large number of experimental observations suggest several additional speculative roles for the CD20 molecule. For example, CD20-specific immunoprecipitation of biochemically cross-linked plasma membrane proteins suggests that CD20 assumes a multimeric structural

conformation characteristic of other previously described membrane channel proteins. Further experimentation has revealed that expression of exogenous CD20 on the cell surface specifically increases Ca^{2+} conductance across the plasma membrane. Together, these results suggest that CD20 complexes may function as B-cell specific Ca^{2+} ion channels. In addition, monoclonal antibodies raised against CD20 have been used to stimulate resting B-cells to transition out of the G0/G1 segment of the cell cycle. It has also been demonstrated that CD20 is associated with both serine and tyrosine kinases and, more specifically, that CD20 is associated, although not directly, with the Src family of tyrosine kinases including p56/53lyn, p56lck, and p59fyn.

A second example of a receptor subfamily, called sialoadhesin molecules, belongs to the Ig superfamily of receptor-like molecules. The more than 100 members of the Ig superfamily are generally considered to engage in specific cell-cell interactions through which intercellular communication may occur. In addition to classical protein-protein interactions, intercellular communication may also be mediated through protein-carbohydrate interactions. In fact, all members of the sialoadhesin family of the Ig superfamily are capable of mediating protein-sialic acid binding interactions. To date, only a small number of proteins have been assigned to the sialoadhesin family including sialoadhesin, CD33, CD22, the myelin-associated glycoprotein (MAG), and the Schwann cell myelin protein (SMP). Each of these proteins is expressed in a restricted subset of cell types. For example, CD22 and CD33 are expressed exclusively by B-lymphocytes and cells of the myelomonocytic lineage, respectively.

Similarly, galectins are a family of the lectin superfamily of carbohydrate-binding proteins which have a high affinity for b-galactoside sugars. Although a large number of glycoproteins containing b-galactoside sugars are produced by the cell, only a few will bind to known galectins *in vitro*. Such apparent binding specificity suggests a highly specific functional role for the galectins. Galectin 1 (conventionally termed *LGALS1* for lectin, galactoside-binding, soluble -1) is thought to specifically bind laminin, a highly polylactosaminated cellular glycoprotein, as well as the highly polylactosaminated lysosome-associated membrane proteins (LAMPs). Galectin 1 has also been shown to bind specifically to a lactosamine-containing glycolipid found on olfactory neurons and to integrin a_7b_1 on skeletal muscle cells. Galectin 3 has also been observed to bind specifically to laminin, immunoglobulin E and its receptor, and bacterial lipopolysaccharides.

Various galectins have been shown to function in the mechanisms of intercellular communication. For example, depending on cell type, galectin 1 has been observed to modulate cell adhesion either positively or negatively. More specifically, galectin 1 appears to inhibit cell adhesion of skeletal muscle presumably by galectin 1-mediated disruption of laminin-integrin a_7b_1 interactions. Alternatively, galectin 1 appears to promote cell adhesion in several non-skeletal muscle cell types examined

presumably by a glycoconjugate cross-linking mechanism. Galectin 3 has also been observed to function in modulating cell-adhesion, as well as in the activation of certain immune cells by cross-linking IgE and IgE receptors. In addition, galectins have been observed to be involved in the regulation of immune cell activity, as well as in such diverse processes as cell adhesion, proliferation, inflammation, autoimmunity, and metastasis of tumor cells. Furthermore, a galectin-like antigen designated HOM-HD-21 was recently found to be highly expressed in a Hodgkin's Disease cDNA library. Very recently, a novel galectin, termed PCTA-1, was identified as a specific cell surface marker on human prostate cancer cell lines and patient-derived carcinomas. Galectins have also been found to function intracellularly as a component of ribonucleoprotein complexes. Finally, galectins 1 and 3 have each been found to modulate T-cell growth and apoptosis by interaction with CD45 and possibly Bcl2, respectively.

A relatively new family of cell-surface proteins has been identified and termed the Ly6 superfamily. The members of this family include murine and human SCA-2, rat Ly-6 (also termed ThB), human CD59 [also known as protectin or membrane attack complex inhibition factor (MACIF)], and E48 antigen. The determination of an initial functional role for SCA-2 may lie in an analysis of its expression profile with regard to the complex process of hematopoiesis. SCA-2 is highly expressed in early thymic precusor cells. In turn, progeny of the intrathymic precusor population continue to express SCA-2, but only until the point of transition occurs from blast cell to small cell. Further experimental evidence demonstrates that mature thymocytes and peripheral T-cells do not express detectable levels of SCA-2, whereas mature, peripheral B-cells do continue to express SCA-2. As a result, it seems very likely that SCA-2 plays an important role in thymocyte maturation and differentiation. A plausible explanation for this functional hypothesis is that SCA-2 may act as a receptor for a unknown cytokine which regulates thymocyte maturation and differentiation.

In addition, CD59 is a recently identified integral membrane protein which appears to be involved in the regulation of complement. Recent studies show that the CD59 antigen may prevent damage from complement C5b-9 and protect astrocytes during inflammatory and infectious disorders of the nervous system. Expression of recombinant human CD59 on porcine donor organs have been shown to prevent complement-mediated lysis and activation of endothelial cells that leads to hyperacute rejection. Recently, researchers at Alexion Pharmaceuticals (New Haven, CT) reported on the production of transgenic pigs which expressed human CD59. In these animals, xenogeneic organs were resistant to hyperacute rejection. (Fodor, et al., "Expression of a functional human complement inhibitor in a transgenic pig as a model for the prevention of xenogeneic hyperacute organ rejection," Proc. Natl. Acad. Sci., 91:1153-11157 (1994).) The same company also reported that expression of recombinant transmembrane CD59 in paroxysmal nocturnal hemoglobinuria (PNH) B-

cells confers resistance to human complement. (Rother et al., "Expression of recombinant transmembrane CD59 in paroxysmal nocturnal hemoglobinuria B-cells confers resistance to human complement," Blood, 84:2604-2611 (1994).) PNH is an acquired hematopoietic disorder characterized by complement-mediated hemolytic anemia, pancytopenia, and venous thrombosis. It is thought that retroviral gene therapy with this molecule could provide a treatment for PNH patients.

A final Ly6 superfamily member, the E48 antigen, is involved in intercellular adhesion between keratinocyte cells of the squamous epithelium. Such keratinocytes are attached to adjoining cells by large numbers of desmosomes, which are thought to play a role in the transition of transformed keratinocytes to metastatic tumor cells. Treatment with a monoclonal antibody raised against the E48 antigen has been successful in the eradication of residual, postoperative squamous cell carcinoma cells of the upper aerodigestive tract in several *in vivo* models and, to some degree, in humans. (van Dongen, et al., "Progress in radioimmunotherapy of head and neck cancer," Oncol. Rep. 1:259-264 (1994).) The gene encoding the E48 antigen has been mapped to the q24-qter region of human chromosome 8. Interestingly, a number of human diseases have been mapped to this region of chromosome 8 including Langer-Giedion syndrome, brachio-otorhinolaryngeal syndrome, trichorhinolaryngeal syndrome, and epidermolysis bullosa simplex.

A further example of a receptor family includes the prohibitin receptors. The prohibitin gene product is expressed in a wide variety of tissues and has been implicated as a component of a number of anti-proliferative mechanisms. The prohibitin gene encodes a 30 kD postsynthetically modified polypeptide located primarily in the mitochondria, but also may be associated with the IgM receptor on the B-cell plasma membrane. The protein functionally inhibits DNA synthesis and entry into S phase of the cell cycle by an unknown mechanism. Interestingly, although the prohibitin gene product is hypothesized to be involved in the maintenance of senescence and the prevention of cancer, one study found that, although somatic mutations in the prohibitin gene were present in a small number of breast cancers, no mutations were identified in any other breast, ovary, liver, and lung cancers examined. (Sato *et al.*, Genomics 17:762-764 (1993).) However, the prohibitin gene has been mapped to human chromosome 17q12-21, the same region thought to contain the gene involved in sporadic breast cancer. Furthermore, DNA sequence analysis of the prohibitin gene identified somatic mutation in 4 of 23 cases of sporadic breast cancer examined. Thus, prohibitin family members may be involved in the development of cancer.

Moreover, the EGFR family of plasma membrane proteins are an integral component of normal cellular proliferation and in the pathogenesis of the cancerous state. The family is relatively small and includes the EGFR, c-erbB-2, c-erbB-3, and others. Various cancers are correlated with aberrant expression of one or more of these

genes. A number of ligands have been identified which bind to the EGFR-like receptors listed above including TGF-a, heparin-binding EGF, amphiregulin, criptoregulin, heregulin, and others. A large fraction of adenocarcinomas examined to date, especially those of the breast, colon, and pancreas, are typified by the amplification or overexpression of the c-erbB-2 gene. EGF, or an analogous ligand, initiates the cellular growth factor response by binding to the EGFR, or EGFR-related, receptor. Following the binding event, the receptor molecule dimerizes activating its intracellular tyrosine kinase domain. This event results in the phosphorylation of specific tyrosine residues near the carboxy terminus of the receptor. The diversity of signals able to be transduced through the relatively small number of EGFR-related receptor molecules is amplified considerably by the recent finding that EGFR-like receptor molecules can function when dimerized with other EGFR family members forming heterodimers.

Members of the EGFR-related family of integral membrane proteins have been implicated in the pathogenesis of a number of human disease-states. For example, a mutation in the EGFR itself appears to play an important role in the development of glioblastomas. (Sang et al., J. Neurosurg 82:841-846 (1995).) The EGFR gene is amplified or overexpressed in the majority of primary human glioblastomas. Although not conferring a distinct advantage on cell growth, an increase in EGFR expression was found to confer an increase in the ability of glioma cells to maintain anchorage-independent growth in soft agar especially in response to EGF and retinoic acid. Anchorage-independent growth *in vitro* correlates highly with tumorigenicity *in vivo*, therefore, it is likely that cells which express abnormally high levels of EGFR in human glioblastoma cells may be involved in the high potential for these cells to cause tumors *in vivo*.

Moreover, overexpression or amplification of c-erbB-2 has been reported to be involved in a high number adenocarcinomas, particularly of the breast, colon, and pancreas, and in a small proportion of ovarian carcinomas.

Thus, there is a clear need for identifying and exploiting novel members of the receptor families, such as those described above. Although structurally related, these receptors will likely possess diverse and multifaceted functions in a variety of cell and tissue types. Receptor type molecules should prove useful in target based screens for small molecules and other such pharmacologically valuable factors. Monoclonal antibodies raised against such receptors may prove useful as therapeutics in an anti-tumor, diagnostic, or other capacity. Furthermore, receptors described here may prove useful in an active or passive immunotherapeutical role in patients with cancer or other immunocompromised disease states.

SUMMARY OF THE INVENTION

In one aspect, the invention relates to receptor polypeptides and polynucleotides, as well as the methods for their production. Another aspect of the invention relates to methods for using such receptor polypeptides and polynucleotides. Such uses include the treatment of the specified diseases, among others. In still another aspect, the invention relates to methods to identify agonists and antagonists using the materials provided by the invention, and treating conditions associated with receptor imbalance with the identified compounds. Yet another aspect of the invention relates to diagnostic assays for detecting diseases associated with inappropriate receptor activity or levels.

DESCRIPTION OF THE INVENTION

Definitions

The following definitions are provided to facilitate understanding of certain terms used frequently herein.

"Receptor" refers, among others, to a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:Y, or an allelic variant thereof.

"Receptor Activity" or "Biological Activity of the Receptor" refers to the metabolic or physiologic function of said receptor including similar activities or improved activities or these activities with decreased undesirable side-effects. Also included are antigenic and immunogenic activities of said receptor.

"Receptor gene" refers to a polynucleotide comprising the nucleotide sequence set forth in SEQ ID NO:X or allelic variants thereof and/or their complements.

"SEQ ID NO:X" comprises all or a substantial portion of the polynucleotide encoding each receptor of the invention. The value X for the nucleotide sequence is an integer specified in Table 1. This nucleotide sequence was translated into the receptor polypeptide identified in Table 1 as "SEQ ID NO:Y," where the value of Y for each receptor polypeptide is an integer defined in Table 1.

The invention further provides a composition of matter comprising a nucleic acid molecule which comprises a human cDNA clone identified by a cDNA Clone ID (Identifier) in Table 1, which DNA molecule is contained in the material deposited with the American Type Culture Collection ("ATCC") and given the ATCC Deposit Number shown in Table 1 for that cDNA clone. The ATCC is located at American Type Culture Collection (ATCC), 12301 Park Lawn Drive, Rockville, Maryland 20852, USA. The deposit has been made under the terms of the Budapest Treaty on the international recognition of the deposit of micro-organisms for purposes of patent procedure. The strain will be irrevocably and without restriction or condition released to the public upon the issuance of a patent. The deposit is provided merely as convenience to those of skill in the art and is not an admission that a deposit is required for enablement, such as that

required under 35 U.S.C. §112. The nucleotide sequence of the polynucleotides contained in the deposited material, as well as the amino acid sequence of the polypeptide encoded thereby, are controlling in the event of any conflict with any description of sequences herein.

"Antibodies" as used herein includes polyclonal and monoclonal antibodies, chimeric, single chain, and humanized antibodies, as well as Fab fragments, including the products of an Fab or other immunoglobulin expression library.

"Isolated" means altered "by the hand of man" from the natural state. If an "isolated" composition or substance occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living animal is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein.

"Polynucleotide" generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. "Polynucleotides" include, without limitation single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, "polynucleotide" refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The term polynucleotide also includes DNAs or RNAs containing one or more modified bases and DNAs or RNAs with backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications has been made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically or metabolically modified forms of polynucleotides as typically found in nature, as well as the chemical forms of DNA and RNA characteristic of viruses and cells. "Polynucleotide" also embraces relatively short polynucleotides, often referred to as oligonucleotides.

"Polypeptide" refers to any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres. "Polypeptide" refers to both short chains, commonly referred to as peptides, oligopeptides or oligomers, and to longer chains, generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene-encoded amino acids. "Polypeptides" include amino acid sequences modified either by natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can

occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York, 1993 and Wold, F., Posttranslational Protein Modifications: Perspectives and Prospects, pgs. 1-12 in POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, 1983; Seifter et al., "Analysis for protein modifications and nonprotein cofactors", Meth Enzymol (1990) 182:626-646 and Rattan et al., "Protein Synthesis: Posttranslational Modifications and Aging", Ann NY Acad Sci (1992) 663:48-62.)

"Variant" as the term is used herein, is a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide respectively, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to

occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis.

"Identity" is a measure of the identity of nucleotide sequences or amino acid sequences. In general, the sequences are aligned so that the highest order match is obtained. "Identity" per se has an art-recognized meaning and can be calculated using published techniques. (See, e.g.: COMPUTATIONAL MOLECULAR BIOLOGY, Lesk, A.M., ed., Oxford University Press, New York, 1988; BIOCOMPUTING: INFORMATICS AND GENOME PROJECTS, Smith, D.W., ed., Academic Press, New York, 1993; COMPUTER ANALYSIS OF SEQUENCE DATA, PART I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; SEQUENCE ANALYSIS IN MOLECULAR BIOLOGY, von Heinje, G., Academic Press, 1987; and SEQUENCE ANALYSIS PRIMER, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991.) While there exist a number of methods to measure identity between two polynucleotide or polypeptide sequences, the term "identity" is well known to skilled artisans. (Carillo, H., and Lipton, D., SIAM J Applied Math (1988) 48:1073.) Methods commonly employed to determine identity or similarity between two sequences include, but are not limited to, those disclosed in Guide to Huge Computers, Martin J. Bishop, ed., Academic Press, San Diego, 1994, and Carillo, H., and Lipton, D., SIAM J Applied Math (1988) 48:1073. Methods to determine identity and similarity are codified in computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, GCS program package (Devereux, J., et al., Nucleic Acids Research (1984) 12(1):387), BLASTP, BLASTN, FASTA (Atschul, S.F. et al., J Molec Biol (1990) 215:403.)

As an illustration, by a polynucleotide having a nucleotide sequence having at least, for example, 95% "identity" to a reference nucleotide sequence of SEQ ID NO:X is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence of SEQ ID NO: X. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. These mutations of the reference sequence may occur at the 5 ' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence.

Similarly, by a polypeptide having an amino acid sequence having at least, for example, 95% "identity" to a reference amino acid sequence of SEQ ID NO:Y is intended that the amino acid sequence of the polypeptide is identical to the reference sequence except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the reference amino acid of SEQ ID NO:Y. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

Polypeptides of the Invention

In one aspect, the present invention relates to receptor polypeptides (or receptor proteins). The receptor polypeptides include the polypeptide of SEQ ID NO:Y; as well as polypeptides comprising the amino acid sequence of SEQ ID NO:Y; and polypeptides comprising the amino acid sequence which have at least 80% identity to that of SEQ ID NO:Y over its entire length, and still more preferably at least 90% identity, and even still more preferably at least 95% identity to SEQ ID NO:Y. Furthermore, those with at least 97-99% identity to SEQ ID NO:Y are highly preferred. Also included within receptor polypeptides are polypeptides having the amino acid sequence which have at least 80% identity to the polypeptide having the amino acid sequence of SEQ ID NO:Y over its entire length, and still more preferably at least 90% identity, and even still more preferably at least 95% identity to SEQ ID NO:Y. Furthermore, those with at least 97-99% are highly preferred. Preferably receptor polypeptides exhibit at least one biological activity of the receptor.

The receptor polypeptides may be in the form of the "mature" protein or may be a part of a larger protein such as a fusion protein. It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification such as multiple histidine residues, or an additional sequence for stability during recombinant production.

Fragments of the receptor polypeptides are also included in the invention. A "fragment" is a polypeptide having an amino acid sequence that entirely is the same as part, but not all, of the amino acid sequence of the aforementioned receptor polypeptides. As with receptor polypeptides, fragments may be "free-standing," or comprised within a larger polypeptide of which they form a part or region, most

preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, and 101 to the end of receptor polypeptide. In this context "about" includes the particularly recited ranges larger or smaller by several, 5, 4, 3, 2 or 1 amino acid at either extreme or at both extremes.

Preferred fragments include, for example, truncation polypeptides having the amino acid sequence of receptor polypeptides, except for deletion of a continuous series of residues that includes the amino terminus, or a continuous series of residues that includes the carboxyl terminus or deletion of two continuous series of residues, one including the amino terminus and one including the carboxyl terminus.

Also preferred are fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. The "domains" of each receptor polypeptide are illustrated in the Figures. The Figures compare SEQ ID NO:Y to the closest know homologue. Identical amino acids shared between the two polypeptides are shaded, while conservative amino acid changes are boxed. By examining the regions or amino acids shaded and/or boxed, the skilled artisan can readily identify conserved domains between the two polypeptides. The amino acids sequences of SEQ ID NO:Y falling within these conserved domains are "fragments" and are specifically contemplated by the present invention. Especially preferred is the extracellular domains of a receptor of the invention. Soluble extracellular domains have antagonist activity mediated by competition with a receptor ligand.

Other preferred fragments are biologically active fragments. Biologically active fragments are those that mediate receptor activity, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Also included are those that are antigenic or immunogenic in an animal, especially in a human.

Preferably, all of these polypeptide fragments retain a biological activity of the receptor, including antigenic activity. Variants of the defined sequence and fragments also form part of the present invention. Preferred variants are those that vary from the referents by conservative amino acid substitutions -- i.e., those that substitute a residue with another of like characteristics. Typical such substitutions are among Ala, Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln; and among the basic residues Lys and Arg; or aromatic residues Phe and Tyr. Particularly preferred are variants in which several, 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination.

The receptor polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

Polynucleotides of the Invention

Another aspect of the invention relates to receptor polynucleotides. Receptor polynucleotides include isolated polynucleotides which encode the receptor polypeptides and fragments, and polynucleotides closely related thereto. More specifically, a receptor polynucleotide of the invention includes a polynucleotide comprising the nucleotide sequence contained in SEQ ID NO:X encoding a receptor polypeptide of SEQ ID NO:Y, and polynucleotide having the particular sequence of SEQ ID NO:X.

Receptor polynucleotides further include a polynucleotide comprising a nucleotide sequence that has at least 80% identity over its entire length to a nucleotide sequence encoding the receptor polypeptide of SEQ ID NO:Y, and a polynucleotide comprising a nucleotide sequence that is at least 80% identical to that of SEQ ID NO:X over its entire length. In this regard, polynucleotides at least 90% identical are particularly preferred, and those with at least 95% are especially preferred. Furthermore, those with at least 97% are highly preferred and those with at least 98-99% are most highly preferred, with at least 99% being the most preferred. Also included under receptor polynucleotides are a nucleotide sequence which has sufficient identity to a nucleotide sequence contained in SEQ ID NO:X, or contained in the cDNA insert in the plasmid deposited with ATCC, to hybridize under conditions useable for amplification or for use as a probe or marker. Moreover, the receptor polynucleotide includes a nucleotide sequence having at least 80% identity to a nucleotide sequence encoding the receptor polypeptide expressed by the cDNA insert deposited at the ATCC, and a nucleotide sequence comprising at least 15 contiguous nucleotides of such cDNA insert. In this regard, polynucleotides at least 90% identical are particularly preferred, and those with at least 95% are especially preferred. Furthermore, those with at least 97% are highly preferred and those with at least 98-99% are most highly preferred, with at least 99% being the most preferred. The invention also provides polynucleotides which are complementary to all the above receptor polynucleotides.

The receptors of the invention are structurally related to other proteins of specified receptor families, as shown by the results in the Figures. The cDNA sequence of SEQ ID NO:X encodes a polypeptide as described in Table 1 as SEQ ID NO:Y. Because the receptor polypeptides contain domains similar in structure to other receptor family members, the receptors of the present invention are expected to have,

inter alia, similar biological functions/properties to their homologous polypeptides and polynucleotides, and their utility is obvious to anyone skilled in the art.

Table 1

Clone ID Name	SEQ ID NO:X	SEQ ID NO:Y	ATCC Deposit No.	ATCC Deposit Date	Receptor Family	Homology
HMACR70	1	18	209054 #####	05/16/97 01/21/98	Ig	Sialoadhesin OB-1
HTEDK48			209054	05/16/97	TM4SF	MRC-OX44 PETA-3
1-1849 bp	2					
160-900 bp	3	19				
HTPED39 HPWAE25	4	20	209054 #####	05/16/97 1/21/98	TM4SF	NAG-2 TALLA-1
HTPEF86	5	21	209053	05/16/97	TM4SF	CD20 B1 Antigen
HSBBF02	6	22	209054	05/16/97	TM4SF	TALLA-1
HLTAH80	7	23	97242 209054	08/02/95 05/16/97	TM4SF	TALLA-1
HTPBA27	8	24	97242 209054	08/02/95 05/16/97	TM4SF	NAG-2
HAIDQ59			209054	05/16/97	TM4SF	CD9 Antigen
5' Sequence	9	25				
3' Sequence	10					
HHFEK40	11	26	209054	05/16/97	TM4SF	PETA-3
HGBGV89	12	27	209125 209054	06/09/97 05/16/97	TM4SF	L6H
HUVBB80	13	28	209054	05/16/97	TM4SF	L6
HJACE54	14	29	209053	05/16/97	Lectin	Galectin-3 Galectin-5 Galectin-8
HROAD63	15	30	209053	05/16/97	Ly6	E48 splice variant
HMWGS46	16	31	209053	05/16/97	Prohibitin	BAP-37
HNFGW06	17	32	209053	05/16/97	EGFR	EGFR

The novel full-length cDNA clone designated **HMACR70** may be a member of the sialoadhesin family of the Ig superfamily of receptor-like molecules and a CD33 homologue. HMACR70 contains a 1497 nucleotide cDNA insert encoding a 315 amino acid ORF and was cloned from a GM-CSF-treated human macrophage cDNA library. The only additional cDNA libraries in the HGS database which include this clone are human eosinophils and possibly human gall bladder. A BLAST analysis of the amino acid sequence of HMACR70 demonstrates that this clone exhibits approximately 50% identity and 69% similarity over a 300 amino acids stretch of a gene termed human

differentiation antigen, and 38% identity and 62% similarity of the human myelin-associated glycoprotein precursor CD33 gene.

A more recent BLAST analysis confirms HMACR70's designation as a sialoadhesin family member. HMACR70 is homologous to two recently identified sialoadhesin family members, human OB binding protein (OB) 1 and 2. (See, Genbank Accession No. U71382; see Figure 1.) It is thought that OB-1 and OB-2 may bind leptin. Thus, HMACR70, as a sialoadhesin family member, may act to attenuate or even amplify intercellular routes of communication, including binding to leptin or modulating the activity of immune cells, such as macrophages. Clearly, any diseases affected by these processes could be treated by the polypeptide or fragment of HMACR70.

The full-length nucleotide sequences of ten novel human cDNA clones which potentially belong to the TM4SF superfamily are disclosed in the table above and will be addressed sequentially.

The cDNA clone **HTEDK48** contains a 1849 nucleotide cDNA insert encoding a 245 amino acid ORF that was cloned from a human testes cDNA library. The coding sequence of HTEDK48 (SEQ ID NO: 3) may be fused to other human proteins, such as 3-hydroxyacyl-CoA dehydrogenase. BLAST analysis of the amino acid sequence of HTEDK48 demonstrates that this clone exhibits approximately 30% identity and 51% similarity over a 245 amino acid stretch of the CD82 molecule. Recent studies have shown that CD82 can associate with CD4 or CD8 and deliver costimulatory signals for the TCR/CD3 pathway. CD82 has also been found to be involved in syncytium formation in HTLV-I-infected T-cells. And finally, in a recently published study in which the expression of the CD82 gene by tumors of the lung was examined retrospectively, it was reported that CD82 may be linked to the suppression of tumor metastasis of prostate cancer. The study also reported that decreased CD82 expression may be involved in malignant progression of such cancers. Thus, HTEDK48 may also be involved in the development of cancer.

A more recent BLAST analysis shows that HTEDK48 is homologous the rat leukocyte antigen, MRC OX-44, and the platelet endothelial tetraspan antigen -3 (PETA-3). (See Figure 2X.) MRC OX-44, a member of a new family of cell surface proteins, appears to be involved in growth regulation. (See, Bellacosa, A., et al., "The Rat Leukocyte antigen MRC OX-44 is a Member of a New Family of Cell Surface Proteins which Appear to be Involved in Growth Regulation," Mol. Cell. Bio. 11: 2864-2872 (1991).) Similarly, PETA-3 has been located to platelet endothelial cells, and an anti-PETA-3 antigen monoclonal antibody can stimulate platelet aggregation and mediator release. (See, Fitter, S., "Molecular Cloning of cDNA Encoding a Novel Platelet-Endothelial Cell Tetra-Span Antigen, PETA-3," Blood, 86(4):1348-1355 (1995).) Thus, HTEDK48 may function similar to MRC OX-44 or PETA-3 to affect

growth of blood cells. Administering polypeptides or fragments of HTEDK48 may be an effective treatment of blood disorders.

The cDNA clone **HPWAE25** contains a 1288 nucleotide cDNA insert encoding a 273 amino acid ORF that was cloned from a human pancreas tumor cDNA library, while clone **HTPED39** represents a truncated cDNA sequence. This clone also appears in a number of other cDNA libraries constructed from a variety of human cell and tissue types including keratinocytes, ulcerative colitis, striatum depression, lymph node breast cancer, ovarian cancer, stage B2 prostate cancer, kidney medulla, and others. Northern blot analysis of HLTAH80 also shows expression in a variety of human cell lines including U937, MM96, WM115, and MDAMB231. A BLAST analysis of the amino acid sequence of HTPED39 demonstrates that this clone exhibits approximately 35% identity and 50% similarity over the entire length of the CD37 molecule. The CD37 antigen is expressed on B cells and on a subpopulation of T cells, but not on pre-B or plasma cells. It has been reported that CD37 expression is downregulated in conjunction with B-cell activation, suggesting that CD37 may be involved in the processes which dictate the activation state of the B-cell.

Moreover, HPWAE25 is also homologous to recently identified TM4SF members, NAG-2 and TALLA-1. (See Figure 3.) NAG-2 is thought to complex with integrins and other TM4SF proteins, while TALLA-1 is a highly specific marker of T-cell acute lymphoblastic leukemia and neuroblastoma. (See, Tachibana, I., et al., "NAG-2, A Novel Transmembrane-4 Superfamily (TM4SF) Protein that Complexs with Integrins and Other TM4SF Proteins," J. Biol. Chem., 272:29181-29189 (1997); Takagi, S., "Identification of a Higly Specific Surface Marker of T-cell Acute Lymphoblastic Leukemia and Neuroblastoma as a New Member of the Transmembrane 4 Superfamily," Int. J. Cancer 61(5):706-715 (1995).) Thus, HPWAE25 may be involved the development of cancer, particularly leukemia, lymphoma, and neuroblastoma. HPWAE25 may be used as an effective treatment of these cancers, as well as a diagnostic marker.

A subfamily of TM4SF receptors include CD20 proteins. A CD20-like cDNA clone was obtained from a human pancreas tumor cDNA library and contains a 1236 nucleotide insert which encodes a 250 amino acid ORF. A BLAST analysis of the deduced amino acid sequence of **HTPEF86** exhibits approximately 41% identity and 61% similarity to the CD20 gene, also known as B1 antigen. (See Figure 4.) Expression of this gene is detected in only two additional HGS human cDNA libraries; amygdala depression and 9 week early stage human. Although the precise functional role of CD20 has yet to be determined, it is clear that CD20 plays a key role in the regulation of B-cell activation. Based primarily on sequence identity, the novel CD20-like molecule presented herein may also be involved in cell cycle activation. Potential therapeutic and/or diagnostic applications for HTPEF86 may include such clinical

presentations as juvenile rheumatoid arthritis, Graves' Disease, and a number of B-cell lymphomas or other lymphoid tumors.

The clone **HSBBF02** contains a 1115 nucleotide cDNA insert encoding a 245 amino acid ORF and was cloned from an HSC 172 cell line cDNA library. This clone also appears in a number of other cDNA libraries constructed from a variety of human cell and tissue types including brain amygdala depression, endothelial cells, fetal liver and heart, osteoblasts, testes, and others. A BLAST analysis of the amino acid sequence of HSBBF02 demonstrates that this clone exhibits approximately 64% identity and 80% similarity with the A15 molecule over a 131 amino acid stretch (A15 is composed of 244 amino acids). A more recent BLAST search shows that HSBBF02 is similar to the TALLA-1 protein and may in fact be a closely related family member. (See Figure 5.)

In addition, a second cDNA clone, designated **HLTAH80**, exhibits sequence similarity to the A15 molecule and TALLA-1. (See Figure 6.) This clone contains a 1662 nucleotide cDNA insert encoding a 253 amino acid ORF and was cloned from a human T-cell lymphoma cDNA library. This clone also appears in a number of other cDNA libraries constructed from a variety of human cell and tissue types including B-cell lymphoma, corpus collosum, endometrial tumor, osteosarcoma, testes, and others. Northern blot analysis of HLTAH80 also shows expression in a variety of human tissues including spleen, lymph node, thymus, PBLs, heart, and a particularly strong signal in skeletal muscle and pancreas. A BLAST analysis of the amino acid sequence of HLTAH80 demonstrates that this clone exhibits approximately 35% identity and 55% similarity over the entire length of the A15 molecule.

Since expression of A15 drops to undetectable levels when comparing immature T-cells to peripheral blood lymphocytes, it is thought that A15 may play a role in the development of T-cells. Furthermore, the MXS1(CCG-B7) gene which codes for A15 contains a number of triplet nucleotide repeats which have been associated with neuropsychiatric diseases such as Huntington's chorea, fragile X syndrome, and myotonic dystrophy. In addition, A15 appears to be expressed exclusively on T-cell acute lymphoblastic leukemia cell lines, including several derived from adult T-cell leukemia and those established by immortalization with human T-cell leukemia virus type 1 or Herpesvirus saimiri. Thus, clones HLTAH80 and/or HSBBF02 may also be involved in diseases caused by the expansion of repeats or chromosomal instability.

The cDNA clone **HTPBA27** contains a 1345 nucleotide cDNA insert encoding a 238 amino acid ORF and was cloned from a human tumor pancreas cDNA library. This clone also appears in a number of other cDNA libraries constructed from a variety of human cell and tissue types including cerebellum, breast lymph node, osteosarcoma, adult testes, RS4;11 bone marrow cell line, microvascular endothelial cells, and others. A BLAST analysis of the amino acid sequence of HTPBA27 demonstrates that this

clone exhibits approximately 40% identity and 64% similarity with a glycoprotein termed CD53 over its entire length. CD53 is thought to be involved in thymopoiesis, since rat CD53 can be detected on immature CD4-8-thymocytes and the functionally mature single-positive subset, but cannot be detected on the intermediate CD4+8+ thymocytic subset of cells. The CD53 molecule has also been implicated as a component of signal transduction pathways in B cells, monocytes and granulocytes, rat macrophages; NK, and T cells. Moreover, as illustrated in Figure 7, HTPBA27 was recently confirmed as a TM4SF receptor. (See, Tachibana, I., et al., "NAG-2, A Novel Transmembrane-4 Superfamily (TM4SF) Protein that with Integrins and Other TM4SF Proteins," J. Biol. Chem., 272:29181-29189 (1997).) Calling the HTPBA27 polypeptide NAG-2, this group confirmed HTPBA27's status as a TM4SF receptor by showing that NAG-2 complexes with integrin and other TM4SF receptors. Thus, diseases caused by the failure of HTPBA27 to complex with integrin and other TM4SF receptors can be treated by administering HTPBA27. HTPBA27 can also be used to diagnose these diseases.

The cDNA clone **HAIDQ59** contains cDNA insert encoding a 221 amino acid ORF that was cloned from a human epithelial cell induced with TNFa and INF cDNA library. The 5' end of HAIDQ59 is represented by the SEQ ID NO: 9, while the 3' end is represented by SEQ ID NO: 10. This clone appears in only two additional cDNA libraries in the HGS database. These two libraries were constructed from the human Jurkat T-cell line and human microvascular endothelial cells. A BLAST analysis of the amino acid sequence of HAIDQ59 demonstrates that this clone exhibits approximately 53% identity and 69% similarity over 226 amino acids of the CD9 TM4SF molecule. (See Figure 8.) It has been demonstrated that the CD9 molecule is involved in signal transduction pathways in platelets, as well as in cell adhesion in both platelets and pre-B-cell lines. Intriguingly, a monoclonal antibody (vpg15), which recognizes the feline homologue of CD9, has been shown to block infection by feline immunodeficiency virus (FIV). Furthermore, a recent study shows that cells expressing high levels of CD9 exhibited suppressed cell motility. Thus, HAIDQ59 may also be involved in signal transduction of blood cells.

The cDNA clone **HHFEK40** contains a 936 nucleotide cDNA insert encoding a 252 amino acid ORF and was cloned from a human fetal heart cDNA library. This clone appears once in the human fetal heart cDNA library and possibly in a hemangiopericytoma cDNA library. A BLAST analysis of the amino acid sequence of HHFEK40 demonstrated that this clone exhibits approximately 60% identity and 75% similarity over the entire length of a molecule designated PETA-3. (See Figure 9.) PETA-3 was originally identified as a novel human platelet surface glycoprotein termed gp27. Although PETA-3 is present in low abundance on the platelet surface, an anti-PETA-3 monoclonal antibody can stimulate platelet aggregation and mediator release.

Thus, HHFEK40 may function similar to PETA-3 to affect growth of blood cells. Administering polypeptides or fragments of HHFEK40 may be an effective treatment of blood disorders.

The cDNA clone **HGBGV89** contains a 738 nucleotide cDNA insert encoding a 197 amino acid ORF and was cloned from a human gall bladder cDNA library. The only two additional appearances of this clone in the HGS database are in a normalized fetal liver cDNA library and in a fetal liver/spleen cDNA library. The cDNA clone **HUVBB80** contains a 1071 nucleotide cDNA insert encoding a 201 amino acid ORF and was cloned from a human umbilical vein cDNA library. This clone appears in several additional cDNA libraries in the HGS database including prostate BPH, thyroid, and fetal liver/spleen. BLAST analyses of the amino acid sequences of HGBGV89 and HUVBB80 demonstrate that these clones exhibit approximately 49% identity and 65% similarity and 47% identity and 68% similarity, respectively, over the entire length of a molecule designated L6 surface protein or human tumor-associated antigen L6. (See Figures 10 & 11.) Moreover, another group has confirmed the TM4SF receptor homology of HGBGV89 by describing the protein as a putative transmembrane protein L6H. (See Genbank Accession No 2587054; see Figure 10.) The L6 cell surface antigen is highly expressed on lung, breast, colon, and ovarian carcinomas. Promising results of phase 1 clinical studies have been reported with an anti-L6 monoclonal antibody, or its humanized counterpart, suggesting that the L6 antigen may be an attractive target for monoclonal antibody-based cancer therapy.

In summary, there is a clear need for identifying and exploiting novel members of the TM4SF superfamily such as those described herein. Although structurally related, these factors will likely possess diverse and multifaceted functions in a variety of cell and tissue types. Receptor type molecules, such as the novel potential members of the TM4SF superfamily detailed here, should prove useful in target based screens for small molecules and other such pharmacologically valuable factors. Monoclonal antibodies raised against such factors may prove useful as therapeutics in an anti-tumor, diagnostic, or other capacity. Furthermore, factors such as the nine novel TM4SF superfamily-like molecules described here may prove useful in an active or passive immunotherapeutical role in patients with cancer or other immunocompromised disease states.

Besides TM4SF receptors, receptors from other families are also described. For example, clone **HJACE54**, also called galectin 11, exhibits significant sequence identity to the rat galectin 5, the chicken galectin 3 gene, and the human galectin 8 genes. (See Figure 12.) The galectin 11 cDNA clone contains an 865 nucleotide insert which encodes a 133 amino acid ORF. The clone was obtained from a Jurkat T-cell G1 phase cDNA library. A BLAST analysis of the deduced amino acid sequence of HJACE54 demonstrates approximately 35% identity and 57% similarity to the amino

acid sequence of the rat galectin 5 gene. Expression of galectin 11 is quite limited in the HGS database. In fact, the only two additional ESTs in the HGS database which contain the HJACE54 sequence were found in human neutrophil and human infant adrenal gland cDNA libraries. Northern blot analyses have not been performed to examine expression patterns of the galectin 11 gene.

Various galectins have been shown to function in the mechanisms of intercellular communication. For example, depending on cell type, galectin 1 has been observed to modulate cell adhesion either positively or negatively. More specifically, galectin 1 appears to inhibit cell adhesion of skeletal muscle presumably by galectin 1-mediated disruption of laminin-integrin a_7b_1 interactions. Alternatively, galectin 1 appears to promote cell adhesion in several non-skeletal muscle cell types examined presumably by a glycoconjugate cross-linking mechanism. Galectin 3 has also been observed to function in modulating cell-adhesion, as well as in the activation of certain immune cells by cross-linking IgE and IgE receptors. In addition, galectins have been observed to be involved in the regulation of immune cell activity, as well as in such diverse processes as cell adhesion, proliferation, inflammation, autoimmunity, and metastasis of tumor cells. Furthermore, a galectin-like antigen designated HOM-HD-21 was recently found to be highly expressed in a Hodgkin's Disease cDNA library. Very recently, a novel galectin, termed PCTA-1, was identified as a specific cell surface marker on human prostate cancer cell lines and patient-derived carcinomas. Galectins have also been found to function intracellularly as a component of ribonucleoprotein complexes. Finally, galectins 1 and 3 have each been found to modulate T-cell growth and apoptosis by interaction with CD45 and possibly Bcl2, respectively. As a result, the discovery of a novel galectin, such as that encoded by HJACE54, is likely to be a valuable asset both diagnostically and therapeutically.

Additionally, a full-length nucleotide sequence of a novel human cDNA clone which encodes an apparent splice variant of the previously described human E48 antigen has recently been determined. (See Figure 13.) Clone **HROAD63** contains a 441 nucleotide cDNA which encodes a 70 amino acid polypeptide. This novel clone exhibits significant sequence identity to several members of a relatively new family of cell-surface proteins termed the Ly6 superfamily. These members include murine and human SCA-2, rat Ly-6 (also termed ThB), and human CD59 [also known as protectin or membrane attack complex inhibition factor (MACIF)]. The novel E48 splice variant was obtained from the HGS human stomach cDNA library. The clone is present in only a limited number of other HGS cDNA libraries including kidney cancer, keratinocyte, and tongue. An alignment of the nucleotide sequences of the human E48 and HROAD63 cDNAs demonstrates that the initial 168 and 178 nucleotides of E48 and HROAD63, respectively, are identical, with the exception of an additional 10 nucleotides of sequence at the extreme 5' end of the HROAD63 sequence. The

sequence of the two clones is also identical for an additional 229 nucleotides including the 3' end of the coding sequences and the entire 3' untranslated regions. The only divergence of nucleotide sequence in this region of the clones is the deletion of a single thymidine residue in the 3' UTR of the E48 cDNA. The major difference between the two nucleotide sequences is a 329 nucleotide deletion from the HROAD63 sequence. This deletion causes a shift in the HROAD63 reading frame and encompasses the translational stop signal used in the E48 clone. As a result, the carboxy terminal sequence of HROAD63 is radically altered with regard to that of E48 (as illustrated in Figure 13 by the obvious differences between amino acids 56-128 of E48 and 56-70 of HROAD63 in the amino acid alignment). The clinical presentation of disorders, including abnormal skin and hair phenotypes, may be attributed, at least in part, to a non-functional Ly6 superfamily member such as E48 or HROAD63. HROAD63 may also be involved in blood disorders, as seen with its homologues SCA-2 and CD59.

A novel prohibitin cDNA clone presented herein was originally identified in a human bone marrow cell line (RS4;11) cDNA library. The clone contains a 1066 nucleotide insert which encodes a 299 amino acid polypeptide. BLAST and BestFit analyses of the predicted amino acid sequence of **HMWGS46** demonstrate a highly significant sequence identity to a murine protein termed IgM B-cell receptor associated protein (BAP)-37 (Genbank accession number X78683). The HMWGS46 amino acid sequence exhibits nearly perfect identity and similarity over the entire length of the murine BAP-37 sequence. (See Figure 14.) In addition, the full-length nucleotide sequences of HMWGS46 and BAP-37 exhibit at least 87% identical. The HMWGS46 clone also exhibits approximately 49% sequence identity and 85% sequence similarity to a human gene designated prohibitin. Finally, the HMWGS46 cDNA appears in a substantial number of HGS human cDNA libraries in addition to the bone marrow cell line cDNA library from which it was cloned. Some of the cDNA libraries in which this clone appears include keratinocytes, induced endothelial cells, activated neutrophils, synovial sarcoma, colon carcinoma cell line, Jurkat cell line membrane bound polysomes, epileptic frontal cortex, primary dendritic cells, and a number of others. The novel gene related to prohibitin and BAP-37 may prove quite useful as a diagnostic for tumorigenesis, as well as a target for therapeutic intervention of such an event. Thus, although the precise functional role of the prohibitin family members are less than clear, it is quite likely that such homologues are involved in such complex processes as development, senescence, and tumor suppression. Therefore a novel gene, such as HMWGS46, may prove quite useful as a diagnostic for tumorigenesis, as well as a target for therapeutic intervention of such an event.

A human cDNA clone encoding a novel epidermal growth factor receptor (EGFR)-like molecule is also disclosed. The novel EGFR-like cDNA clone presented herein was originally identified in an activated human neutrophil cDNA library. The

clone contains a 704 nucleotide insert which encodes a 168 amino acid polypeptide. A BLAST analysis of the predicted amino acid sequence of **HNFGW06** demonstrates that this novel clone exhibits approximately 85% identity and 90% similarity to a protein designated epidermal growth factor receptor-related protein [Homo sapiens]. (See Figure 15.) The expression profile of the HNFGW06 clone in the HGS database indicates the existence of a fairly highly restricted expression pattern. In addition to the activated neutrophil library from which this clone was obtained, it also appears in the following HGS human cDNA libraries: synovial sarcoma, smooth muscle, placenta, and possibly primary dendritic cells.

The novel EGFR-like cDNA clone HNFGW06 may lead to a number of exciting possibilities for therapeutic and/or diagnostic treatments or reagents. For example, HNFGW06 may be involved in the onset of human breast cancers as well. In addition, due to the fact that TGF-a acts through binding to the EGFR, it is possible that HNFGW06 may also play a role in a variety of gastric processes including regulation of acid secretion, regulation of mucous cell growth, and protection against ethanol- and aspirin-induced injury to gastric tissues.

GENERATING POLYNUCLEOTIDES

Polynucleotides of the present invention encoding a receptor may be obtained using standard cloning and screening, from a cDNA library derived from mRNA in cells specified in Table 1 using the expressed sequence tag (EST) analysis (Adams, M.D., et al. Science (1991) 252:1651-1656; Adams, M.D. et al., Nature, (1992) 355:632-634; Adams, M.D., et al., Nature (1995) 377 Supp:3-174.) Polynucleotides of the invention can also be obtained from natural sources such as genomic DNA libraries or can be synthesized using well known and commercially available techniques.

The nucleotide sequence encoding a receptor polypeptide of SEQ ID NO:Y may be identical to the polynucleotide encoding SEQ ID NO:Y, or it may be a sequence, which as a result of the redundancy (degeneracy) of the genetic code, also encodes the polypeptide of SEQ ID NO:Y.

When the polynucleotides of the invention are used for the recombinant production of a receptor polypeptide, the polynucleotide may include the coding sequence for the mature polypeptide or a fragment thereof, by itself; the coding sequence for the mature polypeptide or fragment in reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence, or other fusion peptide portions. For example, a marker sequence which facilitates purification of the fused polypeptide can be encoded. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in

Gentz et al., Proc Natl Acad Sci USA (1989) 86:821-824, or is an HA tag. The polynucleotide may also contain non-coding 5' and 3' sequences, such as transcribed, non-translated sequences, splicing and polyadenylation signals, ribosome binding sites and sequences that stabilize mRNA.

Further preferred embodiments are polynucleotides encoding receptor variants comprising the amino acid sequence of receptor polypeptide of Table 1 (SEQ ID NO:Y) in which several, 5-10, 1-5, 1-3, 1-2 or 1 amino acid residues are substituted, deleted or added, in any combination.

The present invention further relates to polynucleotides that hybridize to the herein above-described sequences. In this regard, the present invention especially relates to polynucleotides which hybridize under stringent conditions to the herein above-described polynucleotides. As herein used, the term "stringent conditions" means hybridization will occur only if there is at least 80%, and preferably at least 90%, and more preferably at least 95%, yet even more preferably 97-99% identity between the sequences.

Polynucleotides of the invention, which are identical or sufficiently identical to a nucleotide sequence contained in SEQ ID NO:X or a fragment thereof, or to the cDNA insert in the plasmid deposited at the ATCC, or a fragment thereof, may be used as hybridization probes for cDNA and genomic DNA, to isolate full-length cDNAs and genomic clones encoding the receptor and to isolate cDNA and genomic clones of other genes (including genes encoding homologs and orthologs) that have a high sequence similarity to the receptor gene. Such hybridization techniques are known to those of skill in the art. Typically these nucleotide sequences are 80% identical, preferably 90% identical, more preferably 95% identical to that of the referent. The probes generally will comprise at least 15 nucleotides. Preferably, such probes will have at least 30 nucleotides and may have at least 50 nucleotides. Particularly preferred probes will range between 30 and 50 nucleotides.

In one embodiment, to obtain a polynucleotide encoding the receptor polypeptide, including homologs and orthologs from other species, comprises the steps of screening an appropriate library under stringent hybridization conditions with a labeled probe having the SEQ ID NO:X or a fragment thereof; and isolating full-length cDNA and genomic clones containing said polynucleotide sequence. Such hybridization techniques are well known to those of skill in the art. Stringent hybridization conditions are as defined above or, alternatively, conditions under overnight incubation at 42°C in a solution comprising: 50% formamide, 5xSSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10 % dextran sulfate, and 20 microgram/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

The polynucleotides and polypeptides of the present invention may be employed as research reagents and materials for discovery of treatments and diagnostics to animal and human disease.

Vectors, Host Cells, Expression

The present invention also relates to vectors which comprise a polynucleotide or polynucleotides of the present invention, and host cells which are genetically engineered with vectors of the invention and to the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention.

For recombinant production, host cells can be genetically engineered to incorporate expression systems or portions thereof for polynucleotides of the present invention. Introduction of polynucleotides into host cells can be effected by methods described in many standard laboratory manuals, such as Davis et al., BASIC METHODS IN MOLECULAR BIOLOGY (1986) and Sambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989) such as calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction or infection.

Representative examples of appropriate hosts include bacterial cells, such as streptococci, staphylococci, E. coli, Streptomyces and Bacillus subtilis cells; fungal cells, such as yeast cells and Aspergillus cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, HEK 293 and Bowes melanoma cells; and plant cells.

A great variety of expression systems can be used. Such systems include, among others, chromosomal, episomal and virus-derived systems, e.g., vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression systems may contain control regions that regulate as well as engender expression. Generally, any system or vector suitable to maintain, propagate or express polynucleotides to produce a polypeptide in a host may be used. The appropriate nucleotide sequence may be inserted into an expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook et al., MOLECULAR CLONING, A LABORATORY MANUAL (supra).

For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the desired polypeptide. These signals may be endogenous to the polypeptide or they may be heterologous signals.

If the receptor polypeptide is to be expressed for use in screening assays, generally, it is preferred that the polypeptide be produced at the surface of the cell. In this event, the cells may be harvested prior to use in the screening assay. If the receptor polypeptide is secreted into the medium, the medium can be recovered in order to recover and purify the polypeptide; if produced intracellularly, the cells must first be lysed before the polypeptide is recovered.

Receptor polypeptides can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during isolation and or purification.

Diagnostic Assays

This invention also relates to the use of receptor polynucleotides or polypeptides for use as diagnostic reagents. Detection of a mutated form of the receptor gene associated with a dysfunction will provide a diagnostic tool that can add to or define a diagnosis of a disease or susceptibility to a disease which results from under-expression, over-expression or altered expression of the receptor. Individuals carrying mutations in the receptor gene may be detected at the DNA level by a variety of techniques.

Nucleic acids for diagnosis may be obtained from a subject's cells, such as from blood, urine, saliva, tissue biopsy or autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR or other amplification techniques prior to analysis. RNA or cDNA may also be used in similar fashion. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to labeled receptor nucleotide sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence differences may also be detected by alterations in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing. (See, e.g., Myers et al., Science (1985) 230:1242.) Sequence changes at specific locations may also be revealed by

nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method. (See Cotton et al., Proc Natl Acad Sci USA (1985) 85: 4397-4401.) In another embodiment, an array of oligonucleotides probes comprising receptor nucleotide sequence or fragments thereof can be constructed to conduct efficient screening of e.g., genetic mutations. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability. (See for example: M.Chee et al., Science, Vol 274, pp 610-613 (1996).)

The diagnostic assays offer a process for diagnosing or determining a susceptibility to specific diseases through detection of mutation in the receptor gene by the methods described.

In addition, specific diseases can be diagnosed by methods comprising determining from a sample derived from a subject an abnormally decreased or increased level of receptor polypeptide or receptor mRNA. Decreased or increased expression can be measured at the RNA level using any of the methods well known in the art for the quantitation of polynucleotides, such as, for example, PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods. Assay techniques that can be used to determine levels of a protein in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays.

Thus in another aspect, the present invention relates to a diagnostic kit for a disease or susceptibility to a disease which comprises:

(a) a receptor polynucleotide, preferably the nucleotide sequence of SEQ ID NO:X, or a fragment thereof ;

(b) a nucleotide sequence complementary to that of (a);

(c) a receptor polypeptide, preferably the polypeptide of SEQ ID NO:Y, or a fragment thereof; or

(d) an antibody to a receptor polypeptide, preferably to the polypeptide of SEQ ID NO: Y.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component.

Chromosome Assays

The nucleotide sequences of the present invention are also valuable for chromosome identification. The sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes according to the present invention is an important first step in correlating those sequences with gene associated disease. Once a sequence has been mapped to a precise chromosomal location, the physical position of the

sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

The differences in the cDNA or genomic sequence between affected and unaffected individuals can also be determined. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent of the disease.

Antibodies

The polypeptides of the invention or their fragments or analogs thereof, or cells expressing them can also be used as immunogens to produce antibodies immunospecific for the receptor polypeptides. The term "immunospecific" means that the antibodies have substantially greater affinity for the polypeptides of the invention than their affinity for other related polypeptides in the prior art.

Antibodies generated against the receptor polypeptides can be obtained by administering the polypeptides or epitope-bearing fragments, analogs or cells to an animal, preferably a nonhuman, using routine protocols. For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler, G. and Milstein, C., Nature (1975) 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor et al., Immunology Today (1983) 4:72) and the EBV-hybridoma technique (Cole et al., MONOCLONAL ANTIBODIES AND CANCER THERAPY, pp. 77-96, Alan R. Liss, Inc., 1985).

Techniques for the production of single chain antibodies (U.S. Patent No. 4,946,778) can also be adapted to produce single chain antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms including other mammals, may be used to express humanized antibodies.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or to purify the polypeptides by affinity chromatography. Antibodies against receptor polypeptides may also be employed to treat diseases.

Vaccines

Another aspect of the invention relates to a method for inducing an immunological response in a mammal which comprises inoculating the mammal with a receptor polypeptide, or a fragment thereof, adequate to produce antibody and/or T cell immune response to protect said animal from a disease. Yet another aspect of the invention relates to a method of inducing immunological response in a mammal which

comprises, delivering a receptor polypeptide via a vector directing expression of the receptor polynucleotide in vivo in order to induce such an immunological response to produce antibody to protect said animal from diseases.

Further aspect of the invention relates to an immunological/vaccine formulation (composition) which, when introduced into a mammalian host, induces an immunological response in that mammal to a receptor polypeptide wherein the composition comprises a receptor polypeptide or receptor gene. The vaccine formulation may further comprise a suitable carrier. Since a receptor polypeptide may be broken down in the stomach, it is preferably administered parenterally (including subcutaneous, intramuscular, intravenous, intradermal etc. injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation instonic with the blood of the recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

Screening Assays

The receptor polypeptide of the present invention may be employed in a screening process for compounds which bind the receptor and which activate (agonists) or inhibit activation of (antagonists) the receptor polypeptide of the present invention. Thus, polypeptides of the invention may also be used to assess the binding of small molecule substrates and ligands in, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These substrates and ligands may be natural substrates and ligands or may be structural or functional mimetics. See Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).

The receptor polypeptides are responsible for many biological functions, including many pathologies. Accordingly, it is desirous to find compounds and drugs which stimulate the receptor on the one hand and which can inhibit the function of the receptor on the other hand. In general, agonists are employed for therapeutic and prophylactic purposes for such conditions and diseases. Antagonists may be employed for a variety of therapeutic and prophylactic purposes for such conditions and diseases.

In general, such screening procedures involve producing appropriate cells which express the receptor polypeptide of the present invention on the surface thereof.

Such cells include cells from mammals, yeast, Drosophila or E. coli. Cells expressing the receptor (or cell membrane containing the expressed receptor) are then contacted with a test compound to observe binding, or stimulation or inhibition of a functional response.

The assays may simply test binding of a candidate compound wherein adherence to the cells bearing the receptor is detected by means of a label directly or indirectly associated with the candidate compound or in an assay involving competition with a labeled competitor. Further, these assays may test whether the candidate compound results in a signal generated by activation of the receptor, using detection systems appropriate to the cells bearing the receptor at their surfaces. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist by the presence of the candidate compound is observed.

Further, the assays may simply comprise the steps of mixing a candidate compound with a solution containing a receptor polypeptide to form a mixture, measuring receptor activity in the mixture, and comparing the receptor activity of the mixture to a standard.

The receptor cDNA, protein and antibodies to the protein may also be used to configure assays for detecting the effect of added compounds on the production of receptor mRNA and protein in cells. For example, an ELISA may be constructed for measuring secreted or cell associated levels of receptor protein using monoclonal and polyclonal antibodies by standard methods known in the art, and this can be used to discover agents which may inhibit or enhance the production of the receptor (also called antagonist or agonist, respectively) from suitably manipulated cells or tissues. Standard methods for conducting screening assays are well understood in the art.

Examples of potential receptor antagonists include antibodies or, in some cases, oligonucleotides or proteins which are closely related to the ligand of the receptor, e.g., a fragment of the ligand, or small molecules which bind to the receptor but do not elicit a response, so that the activity of the receptor is prevented.

Thus in another aspect, the present invention relates to a screening kit for identifying agonists, antagonists, ligands, receptors, substrates, enzymes, etc. for receptor polypeptides; or compounds which decrease or enhance the production of receptor, which comprises:

(a) a receptor polypeptide, preferably that of SEQ ID NO:Y;

(b) a recombinant cell expressing a receptor polypeptide, preferably that of SEQ ID NO:Y;

(c) a cell membrane expressing a receptor polypeptide; preferably that of SEQ ID NO: Y; or

(d) antibody to a receptor polypeptide, preferably that of SEQ ID NO: Y.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component.

Prophylactic and Therapeutic Methods

This invention provides methods of treating an abnormal conditions related to both an excess of and insufficient amounts of receptor activity.

If the activity of the receptor is in excess, several approaches are available. One approach comprises administering to a subject an inhibitor compound (antagonist) as described along with a pharmaceutically acceptable carrier in an amount effective to inhibit activation by blocking the binding of ligands to the receptor or by inhibiting a second signal, and thereby alleviating the abnormal condition.

In another approach, soluble forms of the receptor polypeptides still capable of binding the ligand in competition with endogenous receptor may be administered. Typical embodiments of such competitors comprise fragments of the receptor polypeptide.

In still another approach, expression of the gene encoding endogenous receptor can be inhibited using expression blocking techniques. Known such techniques involve the use of antisense sequences, either internally generated or separately administered. (See, for example, O'Connor, J Neurochem (1991) 56:560 in Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Alternatively, oligonucleotides which form triple helices with the gene can be supplied. (See, for example, Lee et al., Nucleic Acids Res (1979) 6:3073; Cooney et al., Science (1988) 241:456; Dervan et al., Science (1991) 251:1360.) These oligomers can be administered per se or the relevant oligomers can be expressed in vivo.

For treating abnormal conditions related to an under-expression of the receptor and its activity, several approaches are also available. One approach comprises administering to a subject a therapeutically effective amount of a compound which activates the receptor, i.e., an agonist as described above, in combination with a pharmaceutically acceptable carrier, to thereby alleviate the abnormal condition. Alternatively, gene therapy may be employed to effect the endogenous production of the receptor by the relevant cells in the subject. For example, a polynucleotide of the invention may be engineered for expression in a replication defective retroviral vector, as discussed above. The retroviral expression construct may then be isolated and introduced into a packaging cell transduced with a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention such that the packaging cell now produces infectious viral particles containing the gene of interest. These producer cells may be administered to a subject for engineering cells in vivo and expression of the polypeptide in vivo. For overview of gene therapy, see Chapter 20, Gene Therapy and

other Molecular Genetic-based Therapeutic Approaches, (and references cited therein) in Human Molecular Genetics, T Strachan and A P Read, BIOS Scientific Publishers Ltd (1996).

Formulation and Administration

Peptides, such as the soluble form of receptor polypeptides, and agonists and antagonist peptides or small molecules, may be formulated in combination with a suitable pharmaceutical carrier. Such formulations comprise a therapeutically effective amount of the polypeptide or compound, and a pharmaceutically acceptable carrier or excipient. Such carriers include but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. Formulation should suit the mode of administration, and is well within the skill of the art. The invention further relates to pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention.

Polypeptides and other compounds of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds. Preferred forms of systemic administration of the pharmaceutical compositions include injection, typically by intravenous injection. Other injection routes, such as subcutaneous, intramuscular, or intraperitoneal, can be used. Alternative means for systemic administration include transmucosal and transdermal administration using penetrants such as bile salts or fusidic acids or other detergents. In addition, if properly formulated in enteric or encapsulated formulations, oral administration may also be possible. Administration of these compounds may also be topical and/or localized, in the form of salves, pastes, gels and the like.

The dosage range required depends on the choice of peptide, the route of administration, the nature of the formulation, the nature of the subject's condition, and the judgment of the attending practitioner. Suitable dosages, however, are in the range of 0.1-100 µg/kg of subject. Wide variations in the needed dosage, however, are to be expected in view of the variety of compounds available and the differing efficiencies of various routes of administration. For example, oral administration would be expected to require higher dosages than administration by intravenous injection. Variations in these dosage levels can be adjusted using standard empirical routines for optimization, as is well understood in the art.

Polypeptides used in treatment can also be generated endogenously in the subject, in treatment modalities often referred to as "gene therapy" as described above. Thus, for example, cells from a subject may be engineered with a polynucleotide, such as a DNA or RNA, to encode a polypeptide ex vivo, and for example, by the use of a retroviral plasmid vector. The cells are then introduced into the subject.

All publications, including but not limited to patents and patent applications, cited in this specification are herein incorporated by reference as if each individual publication were specifically and individually indicated to be incorporated by reference herein as though fully set forth.

HMACR70

```
GCAGTTCCTGAGAGAAGAACCCTGAGGAACAGACGTTCCCTCGCGGCCCTGGCACCTCCAACCCCAGATATGCTGCTGCTGCTGCTGCTG
CGTCAAGGACTCTCTTCTTGGGACTCCTTGTCTGCAAGGGAGCGCCGGGACCGTGGAGGTTGGGGTCTATACGACGACGACGACGACGAC

                                                                                    M  L  L  L  L  L  L

CCCCTGCTCTGGGGGAGGGAGAGGGTGGAAGGACAGAAGAGTAACCGGAAGGATTACTCGCTGACGATGCAGAGTTCCGTGACCGTGCAA
GGGGACGAGACCCCCTCCCTCTCCCACCTTCCTGTCTTCTCATTGGCCTTCCTAATGAGCGACTGCTACGTCTCAAGGCACTGGCACGTT

 P  L  L  W  G  R  E  R  V  E  G  Q  K  S  N  R  K  D  Y  S  L  T  M  Q  S  S  V  T  V  Q

GAGGGCATGTGTGTCCATGTGCGCTGCTCCTTCTCCTACCCAGTGGACAGCCAGACTGACTCTGACCCAGTTCATGGCTACTGGTTCCGG
CTCCCGTACACACAGGTACACGCGACGAGGAAGAGGATGGGTCACCTGTCGGTCTGACTGAGACTGGGTCAAGTACCGATGACCAAGGCC

 E  G  M  C  V  H  V  R  C  S  F  S  Y  P  V  D  S  Q  T  D  S  D  P  V  H  G  Y  W  F  R

GCAGGGAATGATATAAGCTGGAAGGCTCCAGTGGCCACAAACAACCCAGCTTGGGCAGTGCAGGAGGAAACTCGGGACCGATTCCACCTC
CGTCCCTTACTATATTCGACCTTCCGAGGTCACCGGTGTTTGTTGGGTCGAACCCGTCACGTCCTCCTTTGAGCCCTGGCTAAGGTGGAG

 A  G  N  D  I  S  W  K  A  P  V  A  T  N  N  P  A  W  A  V  Q  E  E  T  R  D  R  F  H  L

CTTGGGGACCCACAGACCAAAAATTGCACCCTGAGCATCAGAGATGCCAGAATGAGTGATGCGGGGAGATACTTCTTTCGTATGGAGAAA
GAACCCCTGGGTGTCTGGTTTTTAACGTGGGACTCGTAGTCTCTACGGTCTTACTCACTACGCCCCTCTATGAAGAAAGCATACCTCTTT

 L  G  D  P  Q  T  K  N  C  T  L  S  I  R  D  A  R  M  S  D  A  G  R  Y  F  F  R  M  E  K

GGAAATATAAAATGGAATTATAAATATGACCAGCTCTCTGTGAACGTGACATACCCTCCTCAGAACTTGACTGTGACTGTCTTCCAAGGA
CCTTTATATTTTACCTTAATATTTATACTGGTCGAGAGACACTTGCACTGTATGGGAGGAGTCTTGAACTGACACTGACAGAAGGTTCCT

 G  N  I  K  W  N  Y  K  Y  D  Q  L  S  V  N  V  T  Y  P  P  Q  N  L  T  V  T  V  F  Q  G

GAAGGCACAGCATCCACAGCTCTGGGGAACAGCTCATCTCTTTCAGTCCTAGAGGGCCAGTCTCTGCGCTTGGTCTGTGCTGTTGACAGC
CTTCCGTGTCGTAGGTGTCGAGACCCCTTGTCGAGTAGAGAAAGTCAGGATCTCCCGGTCAGAGACGCGAACCAGACACGACAACTGTCG

 E  G  T  A  S  T  A  L  G  N  S  S  S  L  S  V  L  E  G  Q  S  L  R  L  V  C  A  V  D  S

AATCCCCCTGCCAGGCTGAGCTGGACCTGGAGGAGTCTGACCCTGTACCCCTCACAGCCCTCAAACCCTCTGGTACTGGAGCTGCAAGTG
TTAGGGGGACGGTCCGACTCGACCTGGACCTCCTCAGACTGGGACATGGGGAGTGTCGGGAGTTTGGGAGACCATGACCTCGACGTTCAC

 N  P  P  A  R  L  S  W  T  W  R  S  L  T  L  Y  P  S  Q  P  S  N  P  L  V  L  E  L  Q  V

CACCTGGGGGATGAAGGGGAATTCACCTGTCGAGCTCAGAACTCTCTGGGTTCCCAGCACGTTTCCCTGAACCTCTCCCTGCAACAGGAG
GTGGACCCCCTACTTCCCCTTAAGTGGACAGCTCGAGTCTTGAGAGACCCAAGGGTCGTGCAAAGGGACTTGGAGAGGGACGTTGTCCTC

 H  L  G  D  E  G  E  F  T  C  R  A  Q  N  S  L  G  S  Q  H  V  S  L  N  L  S  L  Q  Q  E
```

HMACR70

```
TACACAGGCAAAATGAAGCCTGTATCAGGAGTGTTGCTGGGGGCGGTCGGGGGAACTGGAGCCACAGCCCTGGTCTTCCTCTCCTTCTGT
ATGTGTCCGTTTTACTTCGGACATAGTCCTCACAACGACCCCCGCCAGCCCCCTTGACCTCGGTGTCGGGACCAGAAGGAGAGGAAGACA

 Y  T  G  K  M  K  P  V  S  G  V  L  L  G  A  V  G  G  T  G  A  T  A  L  V  F  L  S  F  C

GTCATCTTCATTGTAGTGAGGTCCTGCAGGAAGAAATCGGCAAGACCAGCAGCGGACGTGGGAGACATAGGCATGAAGGATGCAAACACC
CAGTAGAAGTAACATCACTCCAGGACGTCCTTCTTTAGCCGTTCTGGTCGTCGCCTGCACCCTCTGTATCCGTACTTCCTACGTTTGTGG

 V  I  F  I  V  V  R  S  C  R  K  K  S  A  R  P  A  A  D  V  G  D  I  G  M  K  D  A  N  T

ATTCAGGGGCTCAGCCTCTCAGGGTAACTGGATGAGTCCTGGGCAGATGATAACCCCCGACACCATGGCCTGGCTGCCCACTCCCTCAGG
TAAGTCCCCGAGTCGGAGAGTCCCATTGACCTACTCAGGACCCGTCTACTATTGGGGGCTGTGGTACCGGACCGACGGGTGAGGGAGTCC

 I  Q  G  L  S  L  S  G  .
```

(SEQ ID NO:18)

```
GGAGGAAAGAGAGATCCCAGTATGCACCCCTCAGCTTTCATAAGGGGGAGCCTCAGGACCTATCCAGGTCAAGAAGCCACCAACAATGAG
CCTCCTTTCTCTCTAGGGTCATACGTGGGGAGTCGAAAGTATTCCCCCTCGGAGTCCTGGATAGGTCCAGTTCTTCGGTGGTTGTTACTC

TACTCAGAGATCAAGATCCCCAAGTAAGAAAATGCAGAGGCTCGGGCTTGTTTGAGGGTTCACGACCCCTCCAGCAAAGGAGTCTGAGGC
ATGAGTCTCTAGTTCTAGGGGTTCATTCTTTTACGTCTCCGAGCCCGAACAAACTCCCAAGTGCTGGGGAGGTCGTTTCCTCAGACTCCG

TGATTCCAGTAGAATTAGCAGCCCTCAATGCTGTGCAACAAGACATCAGAACTTATTCCTCTTGTCTAACTGAAAATGCATGCCTGATGA
ACTAAGGTCATCTTAATCGTCGGGAGTTACGACACGTTGTTCTGTAGTCTTGAATAAGGAGAACAGATTGACTTTTACGTACGGACTACT

CCAAACTCTCCCTTTCCCCATCCAATCGGTCCACACTCCCCGCCCTGGCCTCTGTACCCACCATTCTCCTCTGTACTTCTCTAAGGATGA
GGTTTGAGAGGGAAAGGGGTAGGTTAGCCAGGTGTGAGGGGCGGGACCGGAGACATGGGTGGTAAGAGGAGACATGAAGAGATTCCTACT

CTACTTTAGATTCCGAATATAGTGAGATTGTAACGTGAAAAAAAAAAAAAAAAAAAA
GATGAAATCTAAGGCTTATATCACTCTAACATTGCACTTTTTTTTTTTTTTTTTTTT
```

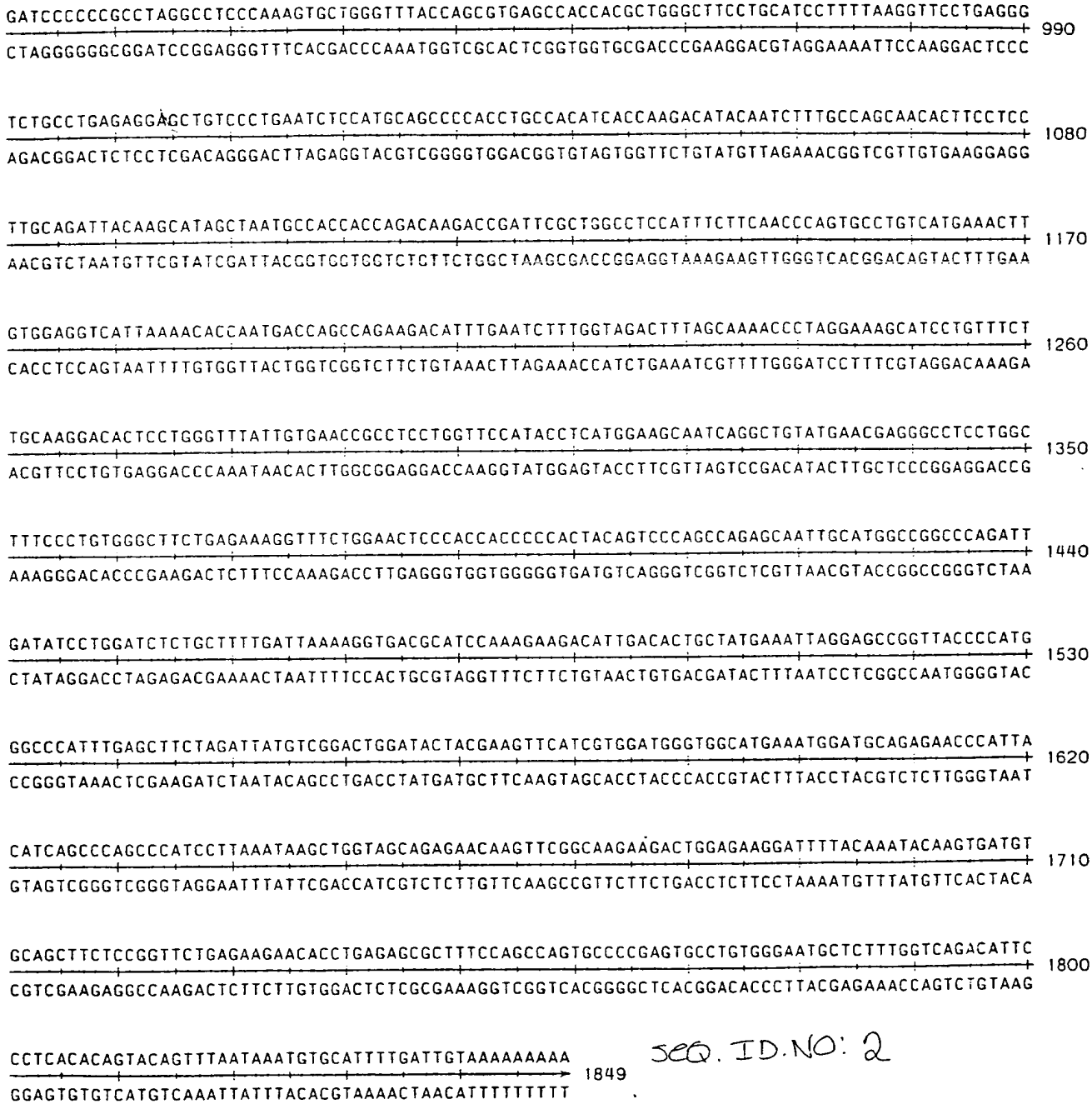
(SEQ ID NO:1)

1497

HTEDK48

```
CGGCACGAGTGGACAACCATCAGGGAGCCAGGACACAGAGGGGCAGAGCAAGTCAGCATTGGCGCCCCTTCCTCAGATCCCTATCATCTT
GCCGTGCTCACCTGTTGGTAGTCCCTCGGTCCTGTGTCTCCCCGTCTCGTTCAGTCGTAACCGCGGGGAAGGAGTCTAGGGATAGTAGAA                 90

GGGAAACAGTAGCCCAGAGGTTCAGGAAGATGTTAACTTAAATGTTCGGGGTGCCCCAGTCTGTTCAGCATGGCTGAAATCCACACTCCG
CCCTTTGTCATCGGGTCTCCAAGTCCTTCTACAATTGAATTTACAAGCCCCACGGGGGTCAGACAAGTCGTACCGACTTTAGGTGTGAGGC                180

                                                                  M  A  E  I  H  T  P

TATTCTTCCTTGAAGAAACTGTTATCTTTACTCAATGGCTTCGTGGCTGTGTCTGGCATCATCCTAGTTGGCCTGGGCATTGGTGGTAAA
ATAAGAAGGAACTTCTTTGACAATAGAAATGAGTTACCGAAGCACCGACACAGACCGTAGTAGGATCAACCGGACCCGTAACCACCATTT                270

 Y  S  S  L  K  K  L  L  S  L  L  N  G  F  V  A  V  S  G  I  I  L  V  G  L  G  I  G  G  K

TGTGGAGGGGCCTCTCTGACGAATGTCCTCGGGCTGTCCTCCGCATACCTCCTTCACGTTGGCAACCTGTGCCTGGTGATGGGATGCATC
ACACCTCCCCGGAGAGACTGCTTACAGGAGCCCGACAGGAGGCGTATGGAGGAAGTGCAACCGTTGGACACGGACCACTACCCTACGTAG                360

 C  G  G  A  S  L  T  N  V  L  G  L  S  S  A  Y  L  L  H  V  G  N  L  C  L  V  M  G  C  I

ACGGTACTGCTTGGCTGTGCCGGGTGGTATGGAGCGACTAAAGAGAGCAGAGGCACGCTCTTGTTTTGCATCCTGTCAATGGTTATTGTC
TGCCATGACGAACCGACACGGCCCACCATACCTCGCTGATTTCTCTCGTCTCCGTGCGAGAACAAAACGTAGGACAGTTACCAATAACAG                450

 T  V  L  L  G  C  A  G  W  Y  G  A  T  K  E  S  R  G  T  L  L  F  C  I  L  S  M  V  I  V

CTCATCATGGAAGTTACAGCTGCCACAGTGGTCCTTCTTTTCTTTCCAATTGTTGGAGATGTGGCCTTGGAACACACCTTCGTGACCCTG
GAGTAGTACCTTCAATGTCGACGGTGTCACCAGGAAGAAAAGAAAGGTTAACAACCTCTACACCGGAACCTTGTGTGGAAGCACTGGGAC                540

 L  I  M  E  V  T  A  A  T  V  V  L  L  F  F  P  I  V  G  D  V  A  L  E  H  T  F  V  T  L

AGGAAGAATTACAGAGGTTACAACGAGCCAGACGACTATTCTACACAGTGGAACTTGGTCATGGAGAAGCTAAAGTGCTGTGGGGTGAAT
TCCTTCTTAATGTCTCCAATGTTGCTCGGTCTGCTGATAAGATGTGTCACCTTGAACCAGTACCTCTTCGATTTCACGACACCCCACTTA                630

 R  K  N  Y  R  G  Y  N  E  P  D  D  Y  S  T  Q  W  N  L  V  M  E  K  L  K  C  C  G  V  N

AACTACACAGATTTTTCTGGCTCTTCCTTCGAAATGACAACGGGCCACACCTACCCCAGGAGTTGCTGTAAATCCATCGGAAGTGTGTCC
TTGATGTGTCTAAAAAGACCGAGAAGGAAGCTTTACTGTTGCCCGGTGTGGATGGGGTCCTCAACGACATTTAGGTAGCCTTCACACAGG                720

 N  Y  T  D  F  S  G  S  S  F  E  M  T  T  G  H  T  Y  P  R  S  C  C  K  S  I  G  S  V  S

TGTGACGGACGCGATGTGTCTCCAAACGTCATCCACCAGAAGGGCTGTTTCCATAAACTCCTAAAAATCACCAAGACTCAGAGCTTCACC
ACACTGCCTGCGCTACACAGAGGTTTGCAGTAGGTGGTCTTCCCGACAAAGGTATTTGAGGATTTTTAGTGGTTCTGAGTCTCGAAGTGG                810

 C  D  G  R  D  V  S  P  N  V  I  H  Q  K  G  C  F  H  K  L  L  K  I  T  K  T  Q  S  F  T

CTGAGTGGGAGCTCTCTGGGAGCTGCAGTGATACAGAGGTGGGGGTCTCGCTATGTTGCGCAGGCTGGTCTTGAACTGCTGGCCTAAAGC   (SEQ ID NO:3)
GACTCACCCTCGAGAGACCCTCGACGTCACTATGTCTCCACCCCCAGAGCGATACAACGCGTCCGACCAGAACTTGACGACCGGATTTCG                900

 L  S  G  S  S  L  G  A  A  V  I  Q  R  W  G  S  R  Y  V  A  Q  A  G  L  E  L  L  A   (SEQ ID NO: 19)
```

HTEDK48



SEQ. ID. NO: 2

HPWAE25

```
GGCGTCCCTCTGCCTGCCCACTCAGTGGCAACACCCGGGAGCTGTTTTGTCCTTTGTGGAGCCTCAGCAGTTCCCTCTTTCAGAACTCAC
                                                                                            90
CCGCAGGGAGACGGACGGGTGAGTCACCGTTGTGGGCCCTCGACAAAACAGGAAACACCTCGGAGTCGTCAAGGGAGAAAGTCTTGAGTG

TGCCAAGAGCCCTGAACAGGAGCCACCATGCAGTGCTTCAGCTTCATTAAGACCATGATGATCCTCTTCAATTTGCTCATCTTTCTGTGT
                                                                                           180
ACGGTTCTCGGGACTTGTCCTCGGTGGTACGTCACGAAGTCGAAGTAATTCTGGTACTACTAGGAGAAGTTAAACGAGTAGAAAGACACA

                                M  Q  C  F  S  F  I  K  T  M  M  I  L  F  N  L  L  I  F  L  C

GGTGCAGCCCTGTTGGCAGTGGGCATCTGGGTGTCAATCGATGGGGCATCCTTTCTGAAGATCTTCGGGCCACTGTCGTCCAGTGCCATG
                                                                                           270
CCACGTCGGGACAACCGTCACCCGTAGACCCACAGTTAGCTACCCCGTAGGAAAGACTTCTAGAAGCCCGGTGACAGCAGGTCACGGTAC

 G  A  A  L  L  A  V  G  I  W  V  S  I  D  G  A  S  F  L  K  I  F  G  P  L  S  S  S  A  M

CAGTTTGTCAACGTGGGCTACTTCCTCATCGCAGCCGGCGTTGTGGTCTTTGCTCTTGGTTTCCTGGGCTGCTATGGTGCTAAGACTGAG
                                                                                           360
GTCAAACAGTTGCACCCGATGAAGGAGTAGCGTCGGCCGCAACACCAGAAACGAGAACCAAAGGACCCGACGATACCACGATTCTGACTC

 Q  F  V  N  V  G  Y  F  L  I  A  A  G  V  V  V  F  A  L  G  F  L  G  C  Y  G  A  K  T  E

AGCAAGTGTGCCCTCGTGACGTTCTTCTTCATCCTCCTCCTCATCTTCATTGCTGAGGTTGCAGCTGCTGTGGTCGCCTTGGTGTACACC
                                                                                           450
TCGTTCACACGGGAGCACTGCAAGAAGAAGTAGGAGGAGGAGTAGAAGTAACGACTCCAACGTCGACGACACCAGCGGAACCACATGTGG

 S  K  C  A  L  V  T  F  F  F  I  L  L  L  I  F  I  A  E  V  A  A  A  V  V  A  L  V  Y  T

ACAATGGCTGAGCACTTCCTGACGTTGCTGGTAGTGCCTGCCATCAAGAAAGATTATGGTTCCCAGGAAGACTTCACTCAAGTGTGGAAC
                                                                                           540
TGTTACCGACTCGTGAAGGACTGCAACGACCATCACGGACGGTAGTTCTTTCTAATACCAAGGGTCCTTCTGAAGTGAGTTCACACCTTG

 T  M  A  E  H  F  L  T  L  L  V  V  P  A  I  K  K  D  Y  G  S  Q  E  D  F  T  Q  V  W  N

ACCACCATGAAAGGGCTCAAGTGCTGTGGCTTCACCAACTATACGGATTTTGAGGACTCACCCTACTTCAAAGAGAACAGTGCCTTTCCC
                                                                                           630
TGGTGGTACTTTCCCGAGTTCACGACACCGAAGTGGTTGATATGCCTAAAACTCCTGAGTGGGATGAAGTTTCTCTTGTCACGGAAAGGG

 T  T  M  K  G  L  K  C  C  G  F  T  N  Y  T  D  F  E  D  S  P  Y  F  K  E  N  S  A  F  P

CCATTCTGTTGCAATGACAACGTCACCAACACAGCCAATGAAACCTGCACCAAGCAAAAGGCTCACGACCAAAAAGTAGAGGGTTGCTTC
                                                                                           720
GGTAAGACAACGTTACTGTTGCAGTGGTTGTGTCGGTTACTTTGGACGTGGTTCGTTTTCCGAGTGCTGGTTTTTCATCTCCCAACGAAG

 P  F  C  C  N  D  N  V  T  N  T  A  N  E  T  C  T  K  Q  K  A  H  D  Q  K  V  E  G  C  F

AATCAGCTTTTGTATGACATCCGAACTAATGCAGTCACCGTGGGTGGTGTGGCAGCTGGAATTGGGGGCCTCGAGCTGGCTGCCATGAAT
                                                                                           810
TTAGTCGAAAACATACTGTAGGCTTGATTACGTCAGTGGCACCCACCACACCGTCGACCTTAACCCCCGGAGCTCGACCGACGGTACTTA

 N  Q  L  L  Y  D  I  R  T  N  A  V  T  V  G  G  V  A  A  G  I  G  G  L  E  L  A  A  M  N

TGTGTCCATGTATCTGTACTGCAATCTACAATAAGTCCACTTCTGCCTCTGCCACTACTGCTGCCACATGGGAACTGTGAAGAGGCACCC
                                                                                           900
ACACAGGTACATAGACATGACGTTAGATGTTATTCAGGTGAAGACGGAGACGGTGATGACGACGGTGTACCCTTGACACTTCTCCGTGGG

 C  V  H  V  S  V  L  Q  S  T  I  S  P  L  L  P  L  P  L  L  L  P  H  G  N  C  E  E  A  P
```

TGGCAAGCAGCAGTGATTGGGGGAGGGGACAGGATCTAACAATGTCACTTGGGCCAGAATGGACCTGCCCTTTCTGCTCCAGACTTGGGG 990
ACCGTTCGTCGTCACTAACCCCCTCCCCTGTCCTAGATTGTTACAGTGAACCCGGTCTTACCTGGACGGGAAAGACGAGGTCTGAACCCC

W Q A A V I G G G D R I . (SEQ ID NO: 20)

CTAGATAGGGACCACTCCTTTTAGGCGATGCCTGACTTTCCTTCCATTGGTGGGTGGATGGGTGGGGGGCATTCCAGAGCCTCTAAGGTA 1080
GATCTATCCCTGGTGAGGAAAATCCGCTACGGACTGAAAGGAAGGTAACCACCCACCTACCCACCCCCCGTAAGGTCTCGGAGATTCCAT

GCCAGTTCTGTTGCCCATTCCCCCAGTCTATTAAACCCTTGATATGCCCCCTAGGCCTAGTGGTGATCCCAGTGCTCTACTGGGGGATGA 1170
CGGTCAAGACAACGGGTAAGGGGGTCAGATAATTTGGGAACTATACGGGGGATCCGGATCACCACTAGGGTCACGAGATGACCCCCTACT

GAGAAAGGCATTTTATAGCCTGGGCATAAGTGAAATCAGCAGAGCCTCTGGGTGGATGTGTAGAAGGCACTTCAAAATGCATAAACCTGT 1260
CTCTTTCCGTAAAATATCGGACCCGTATTCACTTTAGTCGTCTCGGAGACCCACCTACACATCTTCCGTGAAGTTTTACGTATTTGGACA

TACAATGTTAAAAAAAAAAAAAAAAAA 1288 (SEQ ID NO: 4)
ATGTTACAATTTTTTTTTTTTTTTTTT

HTPEF86

```
AAAAAAAACAAGGTCCCCACAGCAAAGAAAAGGAATAGGATCAAGAGATACGTGGCTGCTGGCAGAGCAAGCATGAATTCGATGACTTCA 90
TTTTTTTTGTTCCAGGGGTGTCGTTTCTTTTCCTTATCCTAGTTCTCTATGCACCGACGACCGTCTCGTTCGTACTTAAGCTACTGAAGT

M N S M T S

GCAGTTCCGGTGGCCAATTCTGTGTTGGTGGTGGCACCCCACAATGGTTATCCTGTGACCCCAGGAATTATGTCTCACGTGCCCCTGTAT 180
CGTCAAGGCCACCGGTTAAGACACAACCACCACCGTGGGGTGTTACCAATAGGACACTGGGGTCCTTAATACAGAGTGCACGGGGACATA

A V P V A N S V L V V A P H N G Y P V T P G I M S H V P L Y

CCAAACAGCCAGCCGCAAGTCCACCTAGTTCCTGGGAACCCACCTAGTTTGGTGTCGAATGTGAATGGGCAGCCTGTGCAGAAAGCTCTG 270
GGTTTGTCGGTCGGCGTTCAGGTGGATCAAGGACCCTTGGGTGGATCAAACCACAGCTTACACTTACCCGTCGGACACGTCTTTCGAGAC

P N S Q P Q V H L V P G N P P S L V S N V N G Q P V Q K A L

AAAGAAGGCAAAACCTTGGGGGCCATCCAGATCATCATTGGCCTGGCTCACATCGGCCTCGGCTCCATCATGGCGACGGTTCTCGTAGGG 360
TTTCTTCCGTTTTGGAACCCCCGGTAGGTCTAGTAGTAACCGGACCGAGTGTAGCCGGAGCCGAGGTAGTACCGCTGCCAAGAGCATCCC

K E G K T L G A I Q I I I G L A H I G L G S I M A T V L V G

GAATACCTGTCTATTTCATTCTACGGAGGCTTTCCCTTCTGGGGAGGCTTGTGGTTTATCATTTCAGGATCTCTCTCCGTGGCAGCAGAA 450
CTTATGGACAGATAAAGTAAGATGCCTCCGAAAGGGAAGACCCCTCCGAACACCAAATAGTAAAGTCCTAGAGAGAGGCACCGTCGTCTT

E Y L S I S F Y G G F P F W G G L W F I I S G S L S V A A E

AATCAGCCATATTCTTATTGCCTGCTGTCTGGCAGTTTGGGCTTGAACATCGTCAGTGCAATCTGCTCTGCAGTTGGAGTCATACTCTTC 540
TTAGTCGGTATAAGAATAACGGACGACAGACCGTCAAACCCGAACTTGTAGCAGTCACGTTAGACGAGACGTCAACCTCAGTATGAGAAG

N Q P Y S Y C L L S G S L G L N I V S A I C S A V G V I L F

ATCACAGATCTAAGTATTCCCCACCCATATGCCTACCCCGACTATTATCCTTACGCCTGGGGTGTGAACCCTGGAATGGCGATTTCTGGC 630
TAGTGTCTAGATTCATAAGGGGTGGGTATACGGATGGGGCTGATAATAGGAATGCGGACCCCACACTTGGGACCTTACCGCTAAAGACCG

I T D L S I P H P Y A Y P D Y Y P Y A W G V N P G M A I S G

GTGCTGCTGGTCTTCTGCCTCCTGGAGTTTGGCATCGCATGCGCATCTTCCCACTTTGGCTGCCAGTTGGTCTGCTGTCAATCAAGCAAT 720
CACGACGACCAGAAGACGGAGGACCTCAAACCGTAGCGTACGCGTAGAAGGGTGAAACCGACGGTCAACCAGACGACAGTTAGTTCGTTA

V L L V F C L L E F G I A C A S S H F G C Q L V C C Q S S N

GTGAGTGTCATCTATCCAAACATCTATGCAGCAAACCCAGTGATCACCCCAGAACCGGTGACCTCACCACCAAGTTATTCCAGTGAGATC 810
CACTCACAGTAGATAGGTTTGTAGATACGTCGTTTGGGTCACTAGTGGGGTCTTGGCCACTGGAGTGGTGGTTCAATAAGGTCACTCTAG

V S V I Y P N I Y A A N P V I T P E P V T S P P S Y S S E I
```

HTPEF86

CAAGCAAATAAGTAAGGCTACAGATTCTGGAAGCATCTTTCACTGGGACCAAAAGAAGTCCTCCTCCCTTTCTGGGCTTCCATAACCCAG
GTTCGTTTATTCATTCCGATGTCTAAGACCTTCGTAGAAAGTGACCCTGGTTTTCTTCAGGAGGAGGGAAAGACCCGAAGGTATTGGGTC

Q A N K (SEQIDNO:21)

GTCGTTCCTGTTCTGACAGCTGAGGAAACGTCTCTCCCACTGTTTGTACTCTCACCTTCATTCTTCAATTCAGTCTAGGAAACCATGCTG
CAGCAAGGACAAGACTGTCGACTCCTTTGCAGAGAGGGTGACAAACATGAGAGTGGAAGTAAGAAGTTAAGTCAGATCCTTTGGTACGAC

TTTCTCTATCAAGAAGAAGACAGAGATTTTAAACAGATGTTAACCAAGAGGGACTCCCTAGGGCACATGCATCAGCACATATGTGGGCAT
AAAGAGATAGTTCTTCTTCTGTCTCTAAAATTTGTCTACAATTGGTTCTCCCTGAGGGATCCCGTGTACGTAGTCGTGTATACACCCGTA

CCAGCCTCTGGGGCCTTGGCACACCCATTCGTGTGCTCTGCTGCATGTGAGCTTGTGGGTTAGAGGAACAAATATCTAGACATTCAATCT
GGTCGGAGACCCCGGAACCGTGTGGGTAAGCACACGAGACGACGTACACTCGAACACCCAATCTCCTTGTTTATAGATCTGTAAGTTAGA

TCACTCTTTCAATTGTGCATTCATTTAATAAATAGATACTGAGCATTCAAAAAAAAAAAAAAAAAA
AGTGAGAAAGTTAACACGTAAGTAAATTATTTATCTATGACTCGTAAGTTTTTTTTTTTTTTTTT (SEQIDNO:5)

HSBBF02

```
CACGAGCAGGGTCTCGGGCTAGTCATGGCGTCCCCGTCTCGGAGACTGCAGACTAAACCAGTCATTACTTGTTTCAAGAGCGTTCTGCTA
GTGCTCGTCCCAGAGCCCGATCAGTACCGCAGGGGCAGAGCCTCTGACGTCTGATTTGGTCAGTAATGAACAAAGTTCTCGCAAGACGAT

                  M  A  S  P  S  R  R  L  Q  T  K  P  V  I  T  C  F  K  S  V  L  L

ATCTACACTTTTATTTTCTGGATCACTGGCGTTATCCTTCTTGCAGTTGGCATTTGGGGCAAGGTGAGCCTGGAGAATTACTTTTCTCTT
TAGATGTGAAAATAAAAGACCTAGTGACCGCAATAGGAAGAACGTCAACCGTAAACCCCGTTCCACTCGGACCTCTTAATGAAAAGAGAA

 I  Y  T  F  I  F  W  I  T  G  V  I  L  L  A  V  G  I  W  G  K  V  S  L  E  N  Y  F  S  L

TTAAATGAGAAGGCCACCAATGTCCCCTTCGTGCTCATTGCTACTGGTACCGTCATTATTCTTTTGGGCACCTTTGGTTGTTTTGCTACC
AATTTACTCTTCCGGTGGTTACAGGGGAAGCACGAGTAACGATGACCATGGCAGTAATAAGAAAACCCGTGGAAACCAACAAAACGATGG

 L  N  E  K  A  T  N  V  P  F  V  L  I  A  T  G  T  V  I  I  L  L  G  T  F  G  C  F  A  T

TGCCGAGCTTCTGCATGGATGCTAAAACTGTATGCAATGTTTCTGACTCTCGTTTTTTTGGTCGAACTGGTCGCTGCCATCGTAGGATTT
ACGGCTCGAAGACGTACCTACGATTTTGACATACGTTACAAAGACTGAGAGCAAAAAAACCAGCTTGACCAGCGACGGTAGCATCCTAAA

 C  R  A  S  A  W  M  L  K  L  Y  A  M  F  L  T  L  V  F  L  V  E  L  V  A  A  I  V  G  F

GTTTTCAGACATGAGATTAAGAACAGCTTTAAGAATAATTATGAGAAGGCTTTGAAGCAGTATAACTCTACAGGAGATTATAGAAGCCAT
CAAAAGTCTGTACTCTAATTCTTGTCGAAATTCTTATTAATACTCTTCCGAAACTTCGTCATATTGAGATGTCCTCTAATATCTTCGGTA

 V  F  R  H  E  I  K  N  S  F  K  N  N  Y  E  K  A  L  K  Q  Y  N  S  T  G  D  Y  R  S  H

GCAGTAGACAAGATCCAAAATACGTTGCATTGTTGTGGTGTCACCGATTATAGAGATTGGACAGATACTAATTATTACTCAGAAAAAGGA
CGTCATCTGTTCTAGGTTTTATGCAACGTAACAACACCACAGTGGCTAATATCTCTAACCTGTCTATGATTAATAATGAGTCTTTTTCCT

 A  V  D  K  I  Q  N  T  L  H  C  C  G  V  T  D  Y  R  D  W  T  D  T  N  Y  Y  S  E  K  G

TTTCCTAAGAGTTGCTGTAAACTTGAAGATTGTACTCCACAGAGAGATGCAGACAAAGTAAACAATGAAGGTTGTTTTATAAAGGTGATG
AAAGGATTCTCAACGACATTTGAACTTCTAACATGAGGTGTCTCTCTACGTCTGTTTCATTTGTTACTTCCAACAAAATATTTCCACTAC

 F  P  K  S  C  C  K  L  E  D  C  T  P  Q  R  D  A  D  K  V  N  N  E  G  C  F  I  K  V  M

ACCATTATAGAGTCAGAAATGGGAGTCGTTGCAGGAATTTCCTTTGGAGTTGCTTGCTTCCAACTGATTGGAATCTTTCTCGCCTACTGC
TGGTAATATCTCAGTCTTTACCCTCAGCAACGTCCTTAAAGGAAACCTCAACGAACGAAGGTTGACTAACCTTAGAAAGAGCGGATGACG

 T  I  I  E  S  E  M  G  V  V  A  G  I  S  F  G  V  A  C  F  Q  L  I  G  I  F  L  A  Y  C

CTCTCTCGTGCCATAACAAATAACCAGTATGAGATAGTGTAACCCAATGTATCTGTGGGCCTATTCCTCTCTACCTTTAAGGACATTTAG
GAGAGAGCACGGTATTGTTTATTGGTCATACTCTATCACATTGGGTTACATAGACACCCGGATAAGGAGAGATGGAAATTCCTGTAAATC

 L  S  R  A  I  T  N  N  Q  Y  E  I  V  .   (SEQ ID NO: 22)

GGTCCCCCCTGTGAATTAGAAAGTTGCTTGGCTGGAGAACTGACAACACTACTTACTGATAGACCAAAAAACTACACCAGTAGGTTGATT
CCAGGGGGGACACTTAATCTTTCAACGAACCGACCTCTTGACTGTTGTGATGAATGACTATCTGGTTTTTTGATGTGGTCATCCAACTAA
```

HSBBF02

CAATCAAGATGTATGTAGACCTAAAACTACACCAATAGGCTGATTCAATCAAGATCCGTGCTCGCAGTGGGCTGATTCAATCAAGATGTA 990
GTTAGTTCTACATACATCTGGATTTTGATGTGGTTATCCGACTAAGTTAGTTCTAGGCACGAGCGTCACCCGACTAAGTTAGTTCTACAT

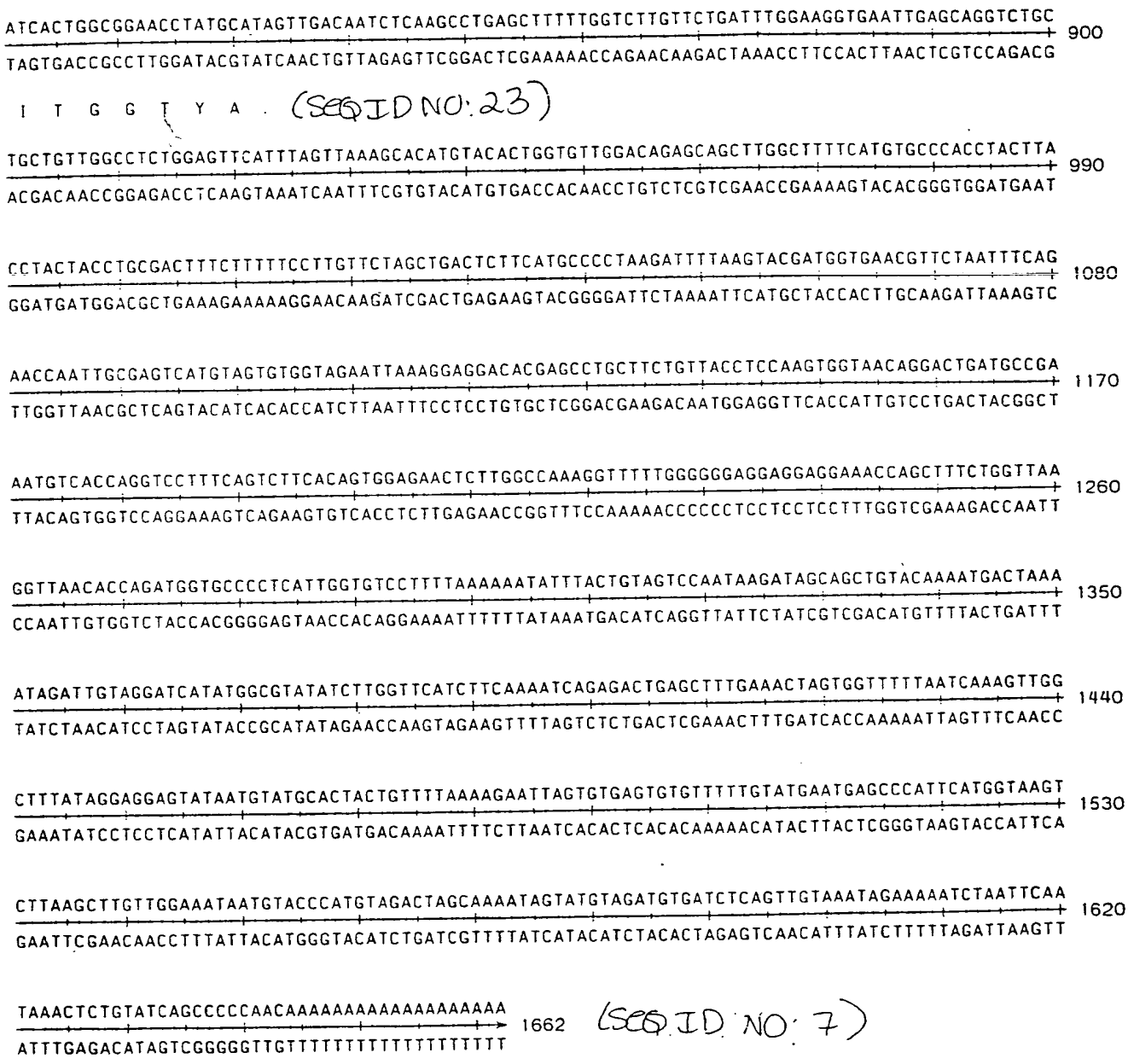
TGTTTGCTATGTTCTAAGTCCACCTTCTATCCCATTCATGTTAGATCGTTGAAACCCTGTATCCCTCTGAAACACTGGAAGAGCTAGTAA 1080
ACAAACGATACAAGATTCAGGTGGAAGATAGGGTAAGTACAATCTAGCAACTTTGGGACATAGGGAGACTTTGTGACCTTCTCGATCATT

ATTGTAAATGAAGTAAAAAAAAAAAAAAAAAAAAA 1115 (SEQ ID NO:6)
TAACATTTACTTCATTTTTTTTTTTTTTTTTTTTT

HLTAH80

```
CACGAGCATTGCCGCTCTCTCGGTGAGCGCAGCCCCGCTCTCCGGGCCGGGCCTTCGCGGGCCACCGGCGCCATGGGCCAGTGCGGCATC
---------+---------+---------+---------+---------+---------+---------+---------+---------+ 90
GTGCTCGTAACGGCGAGAGAGCCACTCGCGTCGGGGCGAGAGGCCCGGCCCGGAAGCGCCCGGTGGCCGCGGTACCCGGTCACGCCGTAG

                                                                          M  G  O  C  G  I

ACCTCCTCCAAGACCGTGCTGGTCTTTCTCAACCTCATCTTCTGGGGGGCAGCTGGCATTTTATGCTATGTGGGAGCCTATGTCTTCATC
---------+---------+---------+---------+---------+---------+---------+---------+---------+ 180
TGGAGGAGGTTCTGGCACGACCAGAAAGAGTTGGAGTAGAAGACCCCCCGTCGACCGTAAAATACGATACACCCTCGGATACAGAAGTAG

 T  S  S  K  T  V  L  V  F  L  N  L  I  F  W  G  A  A  G  I  L  C  Y  V  G  A  Y  V  F  I

ACTTATGATGACTATGACCACTTCTTTGAAGATGTGTACACGCTCATCCCTGCTGTAGTGATCATAGCTGTAGGAGCCCTGCTTTTCATC
---------+---------+---------+---------+---------+---------+---------+---------+---------+ 270
TGAATACTACTGATACTGGTGAAGAAACTTCTACACATGTGCGAGTAGGGACGACATCACTAGTATCGACATCCTCGGGACGAAAAGTAG

 T  Y  D  O  Y  D  H  F  F  E  D  V  Y  T  L  I  P  A  V  V  I  I  A  V  G  A  L  L  F  I

ATTGGGCTAATTGGCTGCTGTGCCACAATCCGGGAAAGTCGCTGTGGACTTGCCACGTTTGTCATCATCCTGCTCTTGGTTTTTGTCACA
---------+---------+---------+---------+---------+---------+---------+---------+---------+ 360
TAACCCGATTAACCGACGACACGGTGTTAGGCCCTTTCAGCGACACCTGAACGGTGCAAACAGTAGTAGGACGAGAACCAAAAACAGTGT

 I  G  L  I  G  C  C  A  T  I  R  E  S  R  C  G  L  A  T  F  V  I  I  L  L  L  V  F  V  T

GAAGTTGTTGTAGTGGTTTTGGGATATGTTTACAGAGCAAAGGTGGAAAATGAGGTTGATCGCAGCATTCAGAAAGTGTATAAGACCTAC
---------+---------+---------+---------+---------+---------+---------+---------+---------+ 450
CTTCAACAACATCACCAAAACCCTATACAAATGTCTCGTTTCCACCTTTTACTCCAACTAGCGTCGTAAGTCTTTCACATATTCTGGATG

 E  V  V  V  V  V  L  G  Y  V  Y  R  A  K  V  E  N  E  V  D  R  S  I  Q  K  V  Y  K  T  Y

AATGGAACCAACCCTGATGCTGCTAGCCGGGCTATTGATTATGTACAGAGACAGCTGCATTGTTGTGGAATTCACAACTACTCAGACTGG
---------+---------+---------+---------+---------+---------+---------+---------+---------+ 540
TTACCTTGGTTGGGACTACGACGATCGGCCCGATAACTAATACATGTCTCTGTCGACGTAACAACACCTTAAGTGTTGATGAGTCTGACC

 N  G  T  N  P  D  A  A  S  R  A  I  D  Y  V  Q  R  Q  L  H  C  C  G  I  H  N  Y  S  D  W

GAAAATACAGATTGGTTCAAAGAAACCAAAAACCAGAGTGTCCCTCTTAGCTGCTGCAGAGAGACTGCCAGCAATTGTAATGGCAGCCTG
---------+---------+---------+---------+---------+---------+---------+---------+---------+ 630
CTTTTATGTCTAACCAAGTTTCTTTGGTTTTTGGTCTCACAGGGAGAATCGACGACGTCTCTCTGACGGTCGTTAACATTACCGTCGGAC

 E  N  T  D  W  F  K  E  T  K  N  Q  S  V  P  L  S  C  C  R  E  T  A  S  N  C  N  G  S  L

GCCCACCCTTCCGACCTCTATGCTGAGGGGTGTGAGGCTCTAGTTGTGAAGAAGCTACAAGAAATCATGATGCATGTGATCTGGGCCGCA
---------+---------+---------+---------+---------+---------+---------+---------+---------+ 720
CGGGTGGGAAGGCTGGAGATACGACTCCCCACACTCCGAGATCAACACTTCTTCGATGTTCTTTAGTACTACGTACACTAGACCCGGCGT

 A  H  P  S  D  L  Y  A  E  G  C  E  A  L  V  V  K  K  L  Q  E  I  M  M  H  V  I  W  A  A

CTGGCATTTGCAGCTATTCAGCTGCTGGGCATGCTGTGTGCTTGCATCGTGTTGTGCAGAAGGAGTAGAGATCCTGCTTACGAGCTCTTC
---------+---------+---------+---------+---------+---------+---------+---------+---------+ 810
GACCGTAAACGTCGATAAGTCGACGACCCGTACGACACACGAACGTAGCACAACACGTCTTCCTCATCTCTAGGACGAATGCTCGAGAAG

 L  A  F  A  A  I  Q  L  L  G  M  L  C  A  C  I  V  L  C  R  R  S  R  D  P  A  Y  E  L  F
```

HLTAH80


ATCACTGGCGGAACCTATGCATAGTTGACAATCTCAAGCCTGAGCTTTTTGGTCTTGTTCTGATTTGGAAGGTGAATTGAGCAGGTCTGC
900
TAGTGACCGCCTTGGATACGTATCAACTGTTAGAGTTCGGACTCGAAAAACCAGAACAAGACTAAACCTTCCACTTAACTCGTCCAGACG
I T G G T Y A . (SEQ ID NO:23)
TGCTGTTGGCCTCTGGAGTTCATTTAGTTAAAGCACATGTACACTGGTGTTGGACAGAGCAGCTTGGCTTTTCATGTGCCCACCTACTTA
990
ACGACAACCGGAGACCTCAAGTAAATCAATTTCGTGTACATGTGACCACAACCTGTCTCGTCGAACCGAAAAGTACACGGGTGGATGAAT
CCTACTACCTGCGACTTTCTTTTTCCTTGTTCTAGCTGACTCTTCATGCCCCTAAGATTTTAAGTACGATGGTGAACGTTCTAATTTCAG
1080
GGATGATGGACGCTGAAAGAAAAAGGAACAAGATCGACTGAGAAGTACGGGGATTCTAAAATTCATGCTACCACTTGCAAGATTAAAGTC
AACCAATTGCGAGTCATGTAGTGTGGTAGAATTAAAGGAGGACACGAGCCTGCTTCTGTTACCTCCAAGTGGTAACAGGACTGATGCCGA
1170
TTGGTTAACGCTCAGTACATCACACCATCTTAATTTCCTCCTGTGCTCGGACGAAGACAATGGAGGTTCACCATTGTCCTGACTACGGCT
AATGTCACCAGGTCCTTTCAGTCTTCACAGTGGAGAACTCTTGGCCAAAGGTTTTTGGGGGGAGGAGGAGGAAACCAGCTTTCTGGTTAA
1260
TTACAGTGGTCCAGGAAAGTCAGAAGTGTCACCTCTTGAGAACCGGTTTCCAAAAACCCCCCTCCTCCTCCTTTGGTCGAAAGACCAATT
GGTTAACACCAGATGGTGCCCCTCATTGGTGTCCTTTTAAAAAATATTTACTGTAGTCCAATAAGATAGCAGCTGTACAAAATGACTAAA
1350
CCAATTGTGGTCTACCACGGGGAGTAACCACAGGAAAATTTTTTATAAATGACATCAGGTTATTCTATCGTCGACATGTTTTACTGATTT
ATAGATTGTAGGATCATATGGCGTATATCTTGGTTCATCTTCAAAATCAGAGACTGAGCTTTGAAACTAGTGGTTTTTAATCAAAGTTGG
1440
TATCTAACATCCTAGTATACCGCATATAGAACCAAGTAGAAGTTTTAGTCTCTGACTCGAAACTTTGATCACCAAAAATTAGTTTCAACC
CTTTATAGGAGGAGTATAATGTATGCACTACTGTTTTAAAAGAATTAGTGTGAGTGTGTTTTTGTATGAATGAGCCCATTCATGGTAAGT
1530
GAAATATCCTCCTCATATTACATACGTGATGACAAAATTTTCTTAATCACACTCACACAAAAACATACTTACTCGGGTAAGTACCATTCA
CTTAAGCTTGTTGGAAATAATGTACCCATGTAGACTAGCAAAATAGTATGTAGATGTGATCTCAGTTGTAAATAGAAAAATCTAATTCAA
1620
GAATTCGAACAACCTTTATTACATGGGTACATCTGATCGTTTTATCATACATCTACACTAGAGTCAACATTTATCTTTTTAGATTAAGTT
TAAACTCTGTATCAGCCCCCAACAAAAAAAAAAAAAAAAAA
1662
ATTTGAGACATAGTCGGGGGTTGTTTTTTTTTTTTTTTTTT
(SEQ ID NO: 7)

HTPBA27

```
CACGAGCGCAGAGCTTGGGGCTTCCTTGGTCGCACCCACCACCTGCCTGCCCACTGGTCAGCCTTCAGGGACCCTGAGCACCGCCTGGTC
                                                                                          90
GTGCTCGCGTCTCGAACCCCGAAGGAACCAGCGTGGGTGGTGGACGGACGGGTGACCAGTCGGAAGTCCCTGGGACTCGTGGCGGACCAG

TCTTTCCTGTGGCCAGCCCAGAACTGAAGCGCTGCGGCATGGCGCGCGCCTGCCTCCAGGCCGTCAAGTACCTCATGTTCGCCTTCAACC
                                                                                         180
AGAAAGGACACCGGTCGGGTCTTGACTTCGCGACGCCGTACCGCGCGCGGACGGAGGTCCGGCAGTTCATGGAGTACAAGCGGAAGTTGG

                                          M  A  R  A  C  L  Q  A  V  K  Y  L  M  F  A  F  N

TGTTCTTCTGGCTGGGAGGCTGTGGCGTGCTGGGTGTCGGCATCTGGCTGGCCGCCACACAGGGGAGCTTCGCCACGCTGTCTTCTTCCT
                                                                                         270
ACAAGAAGACCGACCCTCCGACACCGCACGACCCACAGCCGTAGACCGACCGGCGGTGTGTCCCCTCGAAGCGGTGCGACAGAAGAAGGA

L  F  F  W  L  G  G  C  G  V  L  G  V  G  I  W  L  A  A  T  Q  G  S  F  A  T  L  S  S  S

TCCCGTCCCTGTCGGCTGCCAACCTGCTCATCATCACCGGCGCCTTTGTCATGGCCATCGGCTTCGTGGGCTGCCTGGGTGCCATCAAGG
                                                                                         360
AGGGCAGGGACAGCCGACGGTTGGACGAGTAGTAGTGGCCGCGGAAACAGTACCGGTAGCCGAAGCACCCGACGGACCCACGGTAGTTCC

F  P  S  L  S  A  A  N  L  L  I  I  T  G  A  F  V  M  A  I  G  F  V  G  C  L  G  A  I  K

AGAACAAGTGCCTCCTGCTCACTTTCTTCCTGCTGCTGCTGCTGGTGTTCCTGCTGGAGGCCACCATCGCCATCCTCTTCTTCGCCTACA
                                                                                         450
TCTTGTTCACGGAGGACGAGTGAAAGAAGGACGACGACGACGACCACAAGGACGACCTCCGGTGGTAGCGGTAGGAGAAGAAGCGGATGT

E  N  K  C  L  L  L  T  F  F  L  L  L  L  L  V  F  L  L  E  A  T  I  A  I  L  F  F  A  Y

CGGACAAGATTGACAGGTATGCCCAGCAAGACCTGAAGAAAGGCTTGCACCTGTACGGCACGCAGGGCAACGTGGGCCTCACCAACGCCT
                                                                                         540
GCCTGTTCTAACTGTCCATACGGGTCGTTCTGGACTTCTTTCCGAACGTGGACATGCCGTGCGTCCCGTTGCACCCGGAGTGGTTGCGGA

T  D  K  I  D  R  Y  A  Q  Q  D  L  K  K  G  L  H  L  Y  G  T  Q  G  N  V  G  L  T  N  A

GGAGCATCATCCAGACCGACTTCCGCTGCTGTGGCGTCTCCAACTACACTGACTGGTTCGAGGTGTACAACGCCACGCGGGTACCTGACT
                                                                                         630
CCTCGTAGTAGGTCTGGCTGAAGGCGACGACACCGCAGAGGTTGATGTGACTGACCAAGCTCCACATGTTGCGGTGCGCCCATGGACTGA

W  S  I  I  Q  T  D  F  R  C  C  G  V  S  N  Y  T  D  W  F  E  V  Y  N  A  T  R  V  P  D

CCTGCTGCTTGGAGTTCAGTGAGAGCTGTGGGCTGCACGCCCCCGGCACCTGGTGGAAGGCGCCGTGCTACGAGACGGTGAAGGTGTGGC
                                                                                         720
GGACGACGAACCTCAAGTCACTCTCGACACCCGACGTGCGGGGGCCGTGGACCACCTTCCGCGGCACGATGCTCTGCCACTTCCACACCG

S  C  C  L  E  F  S  E  S  C  G  L  H  A  P  G  T  W  W  K  A  P  C  Y  E  T  V  K  V  W

TTCAGGAGAACCTGCTGGCTGTGGGCATCTTTGGGCTGTGCACGGCGCTGGTGCAGATCCTGGGCCTGACCTTCGCCATGACCATGTACT
                                                                                         810
AAGTCCTCTTGGACGACCGACACCCGTAGAAACCCGACACGTGCCGCGACCACGTCTAGGACCCGGACTGGAAGCGGTACTGGTACATGA

L  Q  E  N  L  L  A  V  G  I  F  G  L  C  T  A  L  V  Q  I  L  G  L  T  F  A  M  T  M  Y

GCCAAGTGGTCAAGGCAGACACCTACTGCGCGTAGGCCGCCCACCGCCGGCTTCTCTGCCAAAAGGACGCCCACGGGGAGATGGCCGCAC
                                                                                         900
CGGTTCACCAGTTCCGTCTGTGGATGACGCGCATCCGGCGGGTGGCGGCCGAAGAGACGGTTTTCCTGCGGGTGCCCCTCTACCGGCGTG

C  Q  V  V  K  A  D  T  Y  C  A  .
```

(SEQ ID NO: 24)

HTPBA27

```
CCACAGCTGCTTTTCCCACCACCAGCTTCGGTGTTCTGCCCCATGCTGGGAGGAGGGAGGGAGGGACAGGTGCCTGGAGCCCCCGGAACC
                                                                                            990
GGTGTCGACGAAAAGGGTGGTGGTCGAAGCCACAAGACGGGGTACGACCCTCCTCCCTCCCTCCCTGTCCACGGACCTCGGGGGCCTTGG

CTGTTTCTGGAAGGCCCTAGCTCAGGTGGCTTCAGGGCCTCCGGACCCCCCCTGGGAGGGGTGGCCACGTGCTGGCTGCGGAACCCAGGG
                                                                                           1080
GACAAAGACCTTCCGGGATCGAGTCCACCGAAGTCCCGGAGGCCTGGGGGGGACCCTCCCCACCGGTGCACGACCGACGCCTTGGGTCCC

CAGGGGTGGGAGGGGCCTCCAGCACTTTTTATATTTACGTATTCTCCAAAGCAGTGTTCACACGGGAGCCAGCCTGTGGCCCCCAGCTTC
                                                                                           1170
GTCCCCACCCTCCCCGGAGGTCGTGAAAAATATAAATGCATAAGAGGTTTCGTCACAAGTGTGCCCTCGGTCGGACACCGGGGGTCGAAG

CTGGAAAACAGGTTGGCGCTGGAGGAGCCGGGTCTTGGCATCCTGGAGGTGGCCCCACTGGTCCTGGTGCTCCAGGCGGGGCCGTGGACC
                                                                                           1260
GACCTTTTGTCCAACCGCGACCTCCTCGGCCCAGAACCGTAGGACCTCCACCGGGGTGACCAGGACCACGAGGTCCGCCCCGGCACCTGG

                                                                          (SEQ ID NO:8)
CCTCACCTACATTCCATAGTGGGCCCGTGGGGCTCCTGGTGCATCTTAATAAAGTGTGAGCAGCAAAAAAAAAAAAAAAAAAAAAAAA
                                                                                           1345
GGAGTGGATGTAAGGTATCACCCGGGCACCCCGAGGACCACGTAGAATTATTTCACACTCGTCGTTTTTTTTTTTTTTTTTTTTTTTTT
```

HAIDQ59

```
GCGCCGCCGGGCCGCAGCATGGGGCGCTTCCGCGGGGGCCTGCGGTGCATCAAGTACCTGCTGCTTGGCTTCAACCTGCTCTTCTGGCTG
CGCGGCGGCCCGGCGTCGTACCCCGCGAAGGCGCCCCCGGACGCCACGTAGTTCATGGACGACGAACCGAAGTTGGACGAGAAGACCGAC
                  M  G  R  F  R  G  G  L  R  C  I  K  Y  L  L  L  G  F  N  L  L  F  W  L

GCTGGATCGGCCGTCATTGCTTTTGGACTATGGTTTCGGTTCGGAGGTGCCATAAAGGAGTTATCATCAGAGGACAAGTCCCCAGAGTAT
CGACCTAGCCGGCAGTAACGAAAACCTGATACCAAAGCCAAGCCTCCACGGTATTTCCTCAATAGTAGTCTCCTGTTCAGGGGTCTCATA
 A  G  S  A  V  I  A  F  G  L  W  F  R  F  G  G  A  I  K  E  L  S  S  E  D  K  S  P  E  Y

TTCTATGTGGGGCTGTATGTTCTGGTTGGAGCCGGGGCCCTGATGATGGCCGTGGGGTTCTTCGGATGCTGCGGAGCCATGCGGGAGTCG
AAGATACACCCCGACATACAAGACCAACCTCGGCCCCGGGACTACTACCGGCACCCCAAGAAGCCTACGACGCCTCGGTACGCCCTCAGC
 F  Y  V  G  L  Y  V  L  V  G  A  G  A  L  M  M  A  V  G  F  F  G  C  C  G  A  M  R  E  S

CAATGTGTGCTTGGATCATTTTTTACCTGCCTCCTGGTGATATTTGCTGCTGAAGTAACCACTGGAGTATTTGCTTTTATAGGCAAGGGG
GTTACACACGAACCTAGTAAAAAATGGACGGAGGACCACTATAAACGACGACTTCATTGGTGACCTCATAAACGAAAATATCCGTTCCCC
 Q  C  V  L  G  S  F  F  T  C  L  L  V  I  F  A  A  E  V  T  T  G  V  F  A  F  I  G  K  G

GTAGCTATCCGACATGTTCAGACCATGTATGAAGAGGCTTACAATGATTACCTTAAAGACAGGGGAAAAGGCAATGGGACACTCATCACC
CATCGATAGGCTGTACAAGTCTGGTACATACTTCTCCGAATGTTACTAATGGAATTTCTGTCCCCTTTTCCGTTACCCTGTGAGTAGTGG
 V  A  I  R  H  V  Q  T  M  Y  E  E  A  Y  N  D  Y  L  K  D  R  G  K  G  N  G  T  L  I  T

TTCCACTCAACATTTCAGTGCTGTGGAAAAGAAAGCTCCGAACAGGTCCAACCTACATGCCCAAAGGAGCTTCTAGGACACAAGAATTGC
AAGGTGAGTTGTAAAGTCACGACACCTTTTCTTTCGAGGCTTGTCCAGGTTGGATGTACGGGTTTCCTCGAAGATCCTGTGTTCTTAACG
 F  H  S  T  F  Q  C  C  G  K  E  S  S  E  Q  V  Q  P  T  C  P  K  E  L  L  G  H  K  N  C

ATCGATGAAATTGAGACCATAATCAGTGTTAAGCTCCAGCTCATTGGAATTGTCGGTATTGGAATTGCAGGTCTGACGATCTTTGGCATG
TAGCTACTTTAACTCTGGTATTAGTCACAATTCGAGGTCGAGTAACCTTAACAGCCATAACCTTAACGTCCAGACTGCTAGAAACCGTAC
 I  D  E  I  E  T  I  I  S  V  K  L  Q  L  I  G  I  V  G  I  G  I  A  G  L  T  I  F  G  M

ATATTCAGCATGGTCCTCTGCTGTGCGATACGAAACTCACGAGATGTGATATGAAGCTACTTCTACATGAAAATTGCAATCTAAAGCTTT
TATAAGTCGTACCAGGAGACGACACGCTATGCTTTGAGTGCTCTACACTATACTTCGATGAAGATGTACTTTTAACGTTAGATTTCGAAA
 I  F  S  M  V  L  C  C  A  I  R  N  S  R  D  V  I  .   (SEQ ID NO:25)

CATACCAAATGTTC
GTATGGTTTACAAG   734   (SEQ ID NO:9)
```

HAIDQ59

AGTGTTTATGGGACTAAAAAACTTTTAACACCTTTTTAGGGGAAATATTTTGGTCCTATACAAAACATGTAAATATGCTTTATTACTTTC 90
TCACAAATACCCTGATTTTTTGAAAATTGTGGAAAAATCCCCTTTATAAAACCAGGATATGTTTTGTACATTTATACGAAATAATGAAAG

ATTTTCTGACCCTGCTGTAAACTACTGCAACCCTCACATCCCTCAAAGGGACTTTTATGTCAAACTCTTCTGTTTCTCCAAATATAAGGA 180
TAAAAGACTGGGACGACATTTGATGACGTTGGGAGTGTAGGGAGTTTCCCTGAAAATACAGTTTGAGAAGACAAAGAGGTTTATATTCCT

AAAAAGACTAAAGCAAGAGATCTGGCAGTTGAAAATTGTGGGAAAGAGAATTTGTATGGGCACTGTATCTATGAAATACCTCATACTTAC 270
TTTTTCTGATTTCGTTCTCTAGACCGTCAACTTTTAACACCCTTTCTCTTAAACATACCCGTGACATAGATACTTTATGGAGTATGAATG

GTTTACATGTTTTCCTAACTTTTTGTATTTTTCTTGTATAGCCACCTAGAGAATTCTTCATAGATTAAGAACTACAGTTTTCACCACTTA 360
CAAATGTACAAAAGGATTGAAAAACATAAAAAGAACATATCGGTGGATCTCTTAAGAAGTATCTAATTCTTGATGTCAAAAGTGGTGAAT

ACATAAGTAAAACAAAGTCCTTCATAATTTAACCATTAGCATCTTTGGCCAAACCAAAATAAAGAAAAGCATCTTCTCCTAGTTGTGTGT 450
TGTATTCATTTTGTTTCAGGAAGTATTAAATTGGTAATCGTAGAAACCGGTTTGGTTTTATTTCTTTTCGTAGAAGAGGATCAACACACA

GGGCAACAGAAACAAGTTAAGGAAACAAAAATACTTATATATACACAGAACAAAAATAATGTTCTTTTTATGCAAATCCCCTGTGAAAAT 540
CCCGTTGTCTTTGTTCAATTCCTTTGTTTTTATGAATATATATGTGTCTTGTTTTTATTACAAGAAAAATACGTTTAGGGGACACTTTTA

AAAATTTTCAATGTTTAAAAAAAAAAAAAAAAAAAA 577 (SEQ ID NO: 10)
TTTTAAAAGTTACAAATTTTTTTTTTTTTTTTTTTT

HHFEK40

```
TTCGGCACGAGCTGCGGGCGGTGGGCGGCTGGGCGGCCCCGGGAGCCGCGCTCTCAGTCTCTCTAGGCGCAGTCCCTTCGCCGCTTCCGG
AAGCCGTGCTCGACGCCCGCCACCCGCCGACCCGCCGGGGCCCTCGGCGCGAGAGTCAGAGAGATCCGCGTCAGGGAAGCGGCGAAGGCC

AGCCCCTGGCAGGGCCCAGAAGCCATGGCCCACTATAAGACTGAGCAGGACGACTGGCTGATCATCTACTTGAAGTATTTACTCTTTGTC
TCGGGGACCGTCCCGGGTCTTCGGTACCGGGTGATATTCTGACTCGTCCTGCTGACCGACTAGTAGATGAACTTCATAAATGAGAAACAG

                    M  A  H  Y  K  T  E  Q  D  D  W  L  I  I  Y  L  K  Y  L  L  F  V

TTCAACTTCTTCTTCTGGGTCGGGGGAGCAGCCGTCCTGGCTGTGGGCATCTGGACCCTGGTGGAGAAGAGTGGCTACCTCAGCGTCCTG
AAGTTGAAGAAGAAGACCCAGCCCCCTCGTCGGCAGGACCGACACCCGTAGACCTGGGACCACCTCTTCTCACCGATGGAGTCGCAGGAC

 F  N  F  F  F  W  V  G  G  A  A  V  L  A  V  G  I  W  T  L  V  E  K  S  G  Y  L  S  V  L

GCCTCCAGCACCTTTGCCGCCTCCGCCTACATCCTCATCTTTGCGGGCGTACTTGTCATGGTGACCGGCTTCCTGGGCTTCGGTGCCATC
CGGAGGTCGTGGAAACGGCGGAGGCGGATGTAGGAGTAGAAACGCCCGCATGAACAGTACCACTGGCCGAAGGACCCGAAGCCACGGTAG

 A  S  S  T  F  A  A  S  A  Y  I  L  I  F  A  G  V  L  V  M  V  T  G  F  L  G  F  G  A  I

CTCTGGGAGCGGAAGGGCTGCCTCTCCACGTATTTCTGCCTGTTGCTCGTCATCTTCCTGGTTGAGCTGGTGGCGGGAGTCCTGGCCCAT
GAGACCCTCGCCTTCCCGACGGAGAGGTGCATAAAGACGGACAACGAGCAGTAGAAGGACCAACTCGACCACCGCCCTCAGGACCGGGTA

 L  W  E  R  K  G  C  L  S  T  Y  F  C  L  L  L  V  I  F  L  V  E  L  V  A  G  V  L  A  H

GTGTATTACCAGAGGCTGAGTGATGAACTGAAGCAGCACTTGAACCGGACTCTGGCTGAGAACTACGGGCAGCCGGAGCACGCAGATCAC
CACATAATGGTCTCCGACTCACTACTTGACTTCGTCGTGAACTTGGCCTGAGACCGACTCTTGATGCCCGTCGGCCTCGTGCGTCTAGTG

 V  Y  Y  Q  R  L  S  D  E  L  K  Q  H  L  N  R  T  L  A  E  N  Y  G  Q  P  E  H  A  D  H

GCCTCAGTGGACCGACTCCAGCAGGATTTCAAGTGCTGCGGAAGCAACAGCTCAGCCGACTGGCAGCACAGCACGTACATCCTGTTGCGG
CGGAGTCACCTGGCTGAGGTCGTCCTAAAGTTCACGACGCCTTCGTTGTCGAGTCGGCTGACCGTCGTGTCGTGCATGTAGGACAACGCC

 A  S  V  D  R  L  Q  Q  D  F  K  C  C  G  S  N  S  S  A  D  W  Q  H  S  T  Y  I  L  L  R

GAGGCCGAGGGCCGCCAGGTGCCCGACAGCTGCTGCAAGACAGTGGTGGCGCGCTGCGGCCAGCGGGCCCACCCCTCCAACATCTATAAG
CTCCGGCTCCCGGCGGTCCACGGGCTGTCGACGACGTTCTGTCACCACCGCGCGACGCCGGTCGCCCGGGTGGGGAGGTTGTAGATATTC

 E  A  E  G  R  Q  V  P  D  S  C  C  K  T  V  V  A  R  C  G  Q  R  A  H  P  S  N  I  Y  K

GTGGAGGGAGGCTGCCTCACCAAGCTGGAGCAGTTCCTGGCCGACCACCTGCTGCTTATGGGGGCAGTGGGCATCGGGGTGGCCTGCCTG
CACCTCCCTCCGACGGAGTGGTTCGACCTCGTCAAGGACCGGCTGGTGGACGACGAATACCCCCGTCACCCGTAGCCCCACCGGACGGAC

 V  E  G  G  C  L  T  K  L  E  Q  F  L  A  D  H  L  L  L  M  G  A  V  G  I  G  V  A  C  L

CAGATCTGCGGGATGGTTCTCACCTGCTGCTTGCACCAGAGGCTCCAGCGGCATTTTTACTAATGGCAACCACCTCCTCTTCCAACTGCC
GTCTAGACGCCCTACCAAGAGTGGACGACGAACGTGGTCTCCGAGGTCGCCGTAAAAATGATTACCGTTGGTGGAGGAGAAGGTTGACGG

 Q  I  C  G  M  V  L  T  C  C  L  H  Q  R  L  Q  R  H  F  Y .
```

(SEQ ID NO:26)

CCTCAAGACAACATGTGGCACATGCCATCTGCAAGG
GGAGTTCTGTTGTACACCGTGTACGGTAGACGTTCC 936 (SEQ ID NO: 11)

HGBGV89

AGCTTACTTTCACTCACCGCCTGTCCTTCCTGACACCTCACCATGTGTACGGGAAAATGTGCCCGCTGTGTGGGGCTCTCCCTCATTACC 90
TCGAATGAAAGTGAGTGGCGGACAGGAAGGACTGTGGAGTGGTACACATGCCCTTTTACACGGGCGACACACCCCGAGAGGGAGTAATGG

M C T G K C A R C V G L S L I T

CTCTGCCTCGTCTGCATTGTGGCCAACGCCCTCCTGCTGGTACCTAATGGGGAGACCTCCTGGACCAACACCAACCATCTCAGCTTGCAA 180
GAGACGGAGCAGACGTAACACCGGTTGCGGGAGGACGACCATGGATTACCCCTCTGGAGGACCTGGTTGTGGTTGGTAGAGTCGAACGTT

L C L V C I V A N A L L L V P N G E T S W T N T N H L S L Q

GTCTGGCTCATGGGCGGCTTCATTGGCGGGGGCCTAATGGTACTGTGTCCAGGGATTGCAGCCGTTCGGGCAGGGGGCAAGGGCTGCTGT 270
CAGACCGAGTACCCGCCGAAGTAACCGCCCCCGGATTACCATGACACAGGTCCCTAACGTCGGCAAGCCCGTCCCCCGTTCCCGACGACA

V W L M G G F I G G G L M V L C P G I A A V R A G G K G C C

GGTGCTGGGTGCTGTGGAAACCGCTGCAGGATGCTGCGCTCGGTCTTCTCCTCGGCGTTCGGGGTGCTTGGTGCCATCTACTG
CCACGACCCACGACACCTTTGGCGACGTCCTACGACGCGAGCCAGAAGAGGAGCCGCAAGCCCCACGAACCACGGTAGATGACGG

G A G C C G N R C R M L R S V F S S A F G V L G A I Y C L

GTGTCTGGAGCTGGGCTCCGAAATGGACCCAGATGCTTAATGAACGGCGAGTGGGGCTACCACTTCGAAGACACCGCGGGAGCTTACTTG 450
CACAGACCTCGACCCGAGGCTTTACCTGGGTCTACGAATTACTTGCCGCTCACCCCGATGGTGAAGCTTCTGTGGCGCCCTCGAATGAAC

V S G A G L R N G P R C L M N G E W G Y H F E D T A G A Y L

CTCAACCGCACTCTATGGGATCGGTGCGAGGCGCCCCCTCGCGTGGTCCCCTGGAATGTGACGCTCTTCTCGCTGCTGGTGGCCGCCTCC 540
GAGTTGGCGTGAGATACCCTAGCCACGCTCCGCGGGGGAGCGCACCAGGGGACCTTACACTGCGAGAAGAGCGACGACCACCGGCGGAGG

L N R T L W D R C E A P P R V V P W N V T L F S L L V A A S

TGCCTGGAGATAGTACTGTGTGGGATCCAGCTGGTGAACGCGACCATTGGTGTCTTCTGCGGCGATTGCAGGAAAAAACAGGACACACCT 630
ACGGACCTCTATCATGACACACCCTAGGTCGACCACTTGCGCTGGTAACCACAGAAGACGCCGCTAACGTCCTTTTTTGTCCTGTGTGGA

C L E I V L C G I Q L V N A T I G V F C G D C R K K Q D T P

CACTGAGGCTCCACTGACCGCCGGGTTACACCTGCTCCTTCCTGGACGCTCACTCCCTTGCTCGCTAGAATAAACTGCTTTGCGCTCTCA 720
GTGACTCCGAGGTGACTGGCGGCCCAATGTGGACGAGGAAGGACCTGCGAGTGAGGGAACGAGCGATCTTATTTGACGAAACGCGAGAGT

H . (SEQ ID NO:27)

AAAAAAAAAAAAAAAAAC 738 (SEQ ID NO:12)
TTTTTTTTTTTTTTTTTG

HUVBR80

```
GGCACGAGAGATTGTCGGCTGCGGGTATATTCCAATTCCCCGTCTCCTCATGAATATGAAGTGAAGGGCTCTGACCCTGGAAGTGGTTCT
                                                                                          90
CCGTGCTCTCTAACAGCCGACGCCCATATAAGGTTAAGGGGCAGAGGAGTACTTATACTTCACTTCCCGAGACTGGGACCTTCACCAAGA

AAGCAGGGCAAAATGGGGTCTCGGAAGTGTGGAGGCTGCCTAAGTTGTTTGCTGATTCCGCTTGCACTTTGGAGTATAATCGTGAACATA
                                                                                         180
TTCGTCCCGTTTTACCCCAGAGCCTTCACACCTCCGACGGATTCAACAAACGACTAAGGCGAACGTGAAACCTCATATTAGCACTTGTAT

            M  G  S  R  K  C  G  G  C  L  S  C  L  L  I  P  L  A  L  W  S  I  I  V  N  I

TTATTGTATTTCCCGAATGGGCAAACTTCCTATGCATCCAGCAATAAACTCACCAACTACGTGTGGTATTTTGAAGGAATCTGTTTCTCA
                                                                                         270
AATAACATAAAGGGCTTACCCGTTTGAAGGATACGTAGGTCGTTATTTGAGTGGTTGATGCACACCATAAAACTTCCTTAGACAAAGAGT

 L  L  Y  F  P  N  G  Q  T  S  Y  A  S  S  N  K  L  T  N  Y  V  W  Y  F  E  G  I  C  F  S

GGCATCATGATGCTTATAGTAACAACAGTTCTTCTGGTACTGGAGAATAATAACAACTATAAATGTTGCCAGAGTGAAAACTGCAGCAAA
                                                                                         360
CCGTAGTACTACGAATATCATTGTTGTCAAGAAGACCATGACCTCTTATTATTGTTGATATTTACAACGGTCTCACTTTTGACGTCGTTT

 G  I  M  M  L  I  V  T  T  V  L  L  V  L  E  N  N  N  N  Y  K  C  C  Q  S  E  N  C  S  K

AAATATGTGACACTGCTGTCAATTATCTTTTCTTCCCTCGGAATTGCTTTTTCTGGATACTGCCTGGTCATCTCTGCCTTGGGTCTTGTC
                                                                                         450
TTTATACACTGTGACGACAGTTAATAGAAAAGAAGGGAGCCTTAACGAAAAAGACCTATGACGGACCAGTAGAGACGGAACCCAGAACAG

 K  Y  V  T  L  L  S  I  I  F  S  S  L  G  I  A  F  S  G  Y  C  L  V  I  S  A  L  G  L  V

CAAGGGCCATATTGCCGCACCCTTGATGGCTGGGAGTATGCTTTTGAAGGCACTGCTGGACGTTTCCTTACAGATTCTAGCATATGGATT
                                                                                         540
GTTCCCGGTATAACGGCGTGGGAACTACCGACCCTCATACGAAAACTTCCGTGACGACCTGCAAAGGAATGTCTAAGATCGTATACCTAA

 Q  G  P  Y  C  R  T  L  D  G  W  E  Y  A  F  E  G  T  A  G  R  F  L  T  D  S  S  I  W  I

CAGTGCCTGGAACCTGCACATGTTGTGGAGTGGAACATCATTTTATTTTCCATTCTCATAACCCTCAGTGGGCTTCAAGTGATCATCTGC
                                                                                         630
GTCACGGACCTTGGACGTGTACAACACCTCACCTTGTAGTAAAATAAAAGGTAAGAGTATTGGGAGTCACCCGAAGTTCACTAGTAGACG

 Q  C  L  E  P  A  H  V  V  E  W  N  I  I  L  F  S  I  L  I  T  L  S  G  L  Q  V  I  I  C

CTCATCAGAGTAGTCATGCAACTATCCAAGATACTGTGTGGAAGCTATTCAGTGATCTTCCAGCCTGGAATCATTTGAATAAGGACAAAA
                                                                                         720
GAGTAGTCTCATCAGTACGTTGATAGGTTCTATGACACACCTTCGATAAGTCACTAGAAGGTCGGACCTTAGTAAACTTATTCCTGTTTT

 L  I  R  V  V  M  Q  L  S  K  I  L  C  G  S  Y  S  V  I  F  Q  P  G  I  I .   (SEQ ID NO:28)

TGTTTTCCATTATCAAGACATGGCCATCTATCTAAATATTATATCAACTGTGTTAGACTTGAGGGCAATATTGAAAATGATGGTGCTTTC
                                                                                         810
ACAAAAGGTAATAGTTCTGTACCGGTAGATAGATTTATAATATAGTTGACACAATCTGAACTCCCGTTATAACTTTTACTACCACGAAAG

TGCATTTGGTGTTTATTTGTAAAAAATTTGCAGTCCTCACTGCACATGCAAGTATACCACCCTTCCATTTAGTATGTTTTTTAAGTAATA
                                                                                         900
ACGTAAACCACAAATAAACATTTTTTAAACGTCAGGAGTGACGTGTACGTTCATATGGTGGGAAGGTAAATCATACAAAAAATTCATTAT
```

HUVBB80

TGCATCAGAAACTTCAGAAATACTTCTGCCCTTTGATCAAACAAATCCATTTCCAAGAATCTGTACTAGGGAAGTAAATAAGAATATGAG 990
ACGTAGTCTTTGAAGTCTTTATGAAGACGGGAAACTAGTTTGTTTAGGTAAAGGTTCTTAGACATGATCCCTTCATTTATTCTTATACTC

(SEQ ID NO:B

AGAAACCTTTATGCAATATGTATATTGCAACATTATTTAATATTCTGGAAAATTGGAAACACCCCAAAATTCTAACTCAAA 1071
TCTTTGGAAATACGTTATACATATAACGTTGTAATAAATTATAAGACCTTTTAACCTTTGTGGGGTTTTAAGATTGAGTTT

HJACE54

```
TTTGTGGAGGGCAGCAGAGAGTACCCAGCTGGACATCCTTTCCTGCTGATGAGCCCCAGGCTGGAGGTGCCCTGCTCACATGCTCTTCCC 90
AAACACCTCCCGTCGTCTCTCATGGGTCGACCTGTAGGAAAGGACGACTACTCGGGGTCCGACCTCCACGGGACGAGTGTACGAGAAGGG

                                                        M  S  P  R  L  E  V  P  C  S  H  A  L  P

CAGGGTCTCTCGCCTGGGCAGGTCATCATAGTACGGGGACTGGTCTTGCAAGAGCCGAAGCATTTTACTGTGAGCCTGAGGGACCAGGCT 180
GTCCCAGAGAGCGGACCCGTCCAGTAGTATCATGCCCCTGACCAGAACGTTCTCGGCTTCGTAAAATGACACTCGGACTCCCTGGTCCGA

 Q  G  L  S  P  G  Q  V  I  I  V  R  G  L  V  L  Q  E  P  K  H  F  T  V  S  L  R  D  Q  A

GCCCATGCTCCTGTGACACTCAGGGCCTCCTTCGCAGACAGAACTCTGGCCTGGATCTCCCGCTGGGGGCAGAAGAAACTGATCTCAGCC 270
CGGGTACGAGGACACTGTGAGTCCCGGAGGAAGCGTCTGTCTTGAGACCGGACCTAGAGGGCGACCCCCGTCTTCTTTGACTAGAGTCGG

 A  H  A  P  V  T  L  R  A  S  F  A  D  R  T  L  A  W  I  S  R  W  G  Q  K  K  L  I  S  A

CCCTTCCTCTTTTACCCCCAGAGATTCTTTGAGGTGCTGCTCCTGTTCCAGGAGGGAGGGCTGAAGCTGGCGCTCAATGGGCAGGGGCTG 360
GGGAAGGAGAAAATGGGGGTCTCTAAGAAACTCCACGACGAGGACAAGGTCCTCCCTCCCGACTTCGACCGCGAGTTACCCGTCCCCGAC

 P  F  L  F  Y  P  Q  R  F  F  E  V  L  L  L  F  Q  E  G  G  L  K  L  A  L  N  G  Q  G  L

GGGGCCACCAGCATGAACCAGCAGGCCCTGGAGCAGCTGCGGGAGCTCCGGATCAGTGGAAGTGTCCAGCTCTACTGTGTCCACTCCTGA 450
CCCCGGTGGTCGTACTTGGTCGTCCGGGACCTCGTCGACGCCCTCGAGGCCTAGTCACCTTCACAGGTCGAGATGACACAGGTGAGGACT

 G  A  T  S  M  N  Q  Q  A  L  E  Q  L  R  E  L  R  I  S  G  S  V  Q  L  Y  C  V  H  S  .  (SEQ ID NO:29)

AGGATGGTTCCAGGAAATACCGCAGAAAACAAGAGTCAGCCACTCCCCAGGGCCCCACTCTCCTCCCCTCATTAAACCATCCACCTGAAC 540
TCCTACCAAGGTCCTTTATGGCGTCTTTTGTTCTCAGTCGGTGAGGGGTCCCGGGGTGAGAGGAGGGGAGTAATTTGGTAGGTGGACTTG

ACCAGCACATCAGGGCCTGGTTCACCTCTGGGGTCACGAGACTGAGTCTACAGGAGCTTTGGGCCTGAGGGAAGGCACAAGAGTGCAAAG 630
TGGTCGTGTAGTCCCGGACCAAGTGGAGACCCCAGTGCTCTGACTCAGATGTCCTCGAAACCCGGACTCCCTTCCGTGTTCTCACGTTTC

GTTCCTCGAACTCTGCACCTTCCTCCACCAGGAGCCTGGGATATGGCTCCATCTGCCTTCAGGGCCTGGACTGCACTCACAGAGGCAAGT 720
CAAGGAGCTTGAGACGTGGAAGGAGGTGGTCCTCGGACCCTATACCGAGGTAGACGGAAGTCCCGGACCTGACGTGAGTGTCTCCGTTCA

GTTGTAGACTAACAAAGATACTCCAAAATACAATGGCTTAAAGAATGTGGTCATTTATTCTTTATTATTTATTTATTTGTGGTCAAATAA 810
CAACATCTGATTGTTTCTATGAGGTTTTATGTTACCGAATTTCTTACACCAGTAAATAAGAAATAATAAATAAATAAACACCAGTTTATT

ATAAATAAGGTTATTTATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 865 (SEQ ID NO: 14)
TATTTATTCCAATAAATAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
```

HROAD63

```
GCACGAGAGACGACATCAGAGATGAGGACAGCATTGCTGCTCCTTGCAGCCCTGGCTGTGGCTACAGGGCCAGCCCTTACCCTGCGCTGC
                                                                                          90
CGTGCTCTCTGCTGTAGTCTCTACTCCTGTCGTAACGACGAGGAACGTCGGGACCGACACCGATGTCCCGGTCGGGAATGGGACGCGACG

                  M  R  T  A  L  L  L  L  A  A  L  A  V  A  T  G  P  A  L  T  L  R  C

CACGTGTGCACCAGCTCCAGCAACTGCAAGCATTCTGTGGTCTGCCCGGCCAGCTCTCGCTTCTGCAAGACCACGAACACAGTGGAGCCT
                                                                                         180
GTGCACACGTGGTCGAGGTCGTTGACGTTCGTAAGACACCAGACGGGCCGGTCGAGAGCGAAGACGTTCTGGTGCTTGTGTCACCTCGGA

 H  V  C  T  S  S  S  N  C  K  H  S  V  V  C  P  A  S  S  R  F  C  K  T  T  N  T  V  E  P

CTGAGGGCTTCCCCGAAAGTCTGGGACCAGGTCCAGGTGGGCATGGAATGCTGATGACTTGGAGCAGGCCCCACAGACCCCACAGAGGAT
                                                                                         270
GACTCCCGAAGGGGCTTTCAGACCCTGGTCCAGGTCCACCCGTACCTTACGACTACTGAACCTCGTCCGGGGTGTCTGGGGTGTCTCCTA

 L  R  A  S  P  K  V  W  D  Q  V  Q  V  G  M  E  C  .
```

(SEQ ID NO: 30)

```
GAAGCCACCCCACAGAGGATGCAGCCCCCAGCTGCATGGAAGGTGGAGGACAGAAGCCCTGTGGATCCCCGGATTTCACACTCCTTCTGT
                                                                                         360
CTTCGGTGGGGTGTCTCCTACGTCGGGGGTCGACGTACCTTCCACCTCCTGTCTTCGGGACACCTAGGGGCCTAAAGTGTGAGGAAGACA

TTTGTTGCCGTTTATTTTTGTACTCAAATCTCTACATGGAGATAAATGATTTAAACCAGTAAAAAAAAAAAAAAAAAAAA
                                                                                    441
AAACAACGGCAAATAAAAACATGAGTTTAGAGATGTACCTCTATTTACTAAATTTGGTCATTTTTTTTTTTTTTTTTTTT
```

(SEQ ID NO: 15)

```
AGCGGGCCCGAACCCTCGTGTGAAGGGTGCAGTACCTAAGCCGGAGCGGGGTAGAGGCGGGCCGGCACCCCCTTCTGACCTCCAGTGCCG
TCGCCCGGGCTTGGGAGCACACTTCCCACGTCATGGATTCGGCCTCGCCCCATCTCCGCCCGGCCGTGGGGGAAGACTGGAGGTCACGGC                  90

CCGGCCTCAAGATCAGACATGGCCCAGAACTTGAAGGACTTGGCGGGACGGCTGCCCGCCGGGCCCCGGGGCATGGGCACGGCCCTGAAG
GGCCGGAGTTCTAGTCTGTACCGGGTCTTGAACTTCCTGAACCGCCCTGCCGACGGGCGGCCCGGGGCCCCGTACCCGTGCCGGGACTTC                 180

                    M  A  Q  N  L  K  D  L  A  G  R  L  P  A  G  P  R  G  M  G  T  A  L  K

CTGTTGCTGGGGGCCGGCGCCGTGGCCTACGGTGTGCGCGAATCTGTGTTCACCGTGGAAGGCGGGCACAGAGCCATCTTCTTCAATCGG
GACAACGACCCCCGGCCGCGGCACCGGATGCCACACGCGCTTAGACACAAGTGGCACCTTCCGCCCGTGTCTCGGTAGAAGAAGTTAGCC                 270

 L  L  L  G  A  G  A  V  A  Y  G  V  R  E  S  V  F  T  V  E  G  G  H  R  A  I  F  F  N  R

ATCGGTGGAGTGCAGCAGGACACTATCCTGGCCGAGGGCCTTCACTTCAGGATCCCTTGGTTCCAGTACCCCATTATCTATGACATTCGG
TAGCCACCTCACGTCGTCCTGTGATAGGACCGGCTCCCGGAAGTGAAGTCCTAGGGAACCAAGGTCATGGGGTAATAGATACTGTAAGCC                 360

 I  G  G  V  Q  Q  D  T  I  L  A  E  G  L  H  F  R  I  P  W  F  Q  Y  P  I  I  Y  D  I  R

GCCAGACCTCGAAAAATCTCCTCCCCTACAGGCTCCAAAGACCTACAGATGGTGAATATCTCCCTGCGAGTGTTGTCTCGACCCAATGCT
CGGTCTGGAGCTTTTTAGAGGAGGGGATGTCCGAGGTTTCTGGATGTCTACCACTTATAGAGGGACGCTCACAACAGAGCTGGGTTACGA                 450

 A  R  P  R  K  I  S  S  P  T  G  S  K  D  L  Q  M  V  N  I  S  L  R  V  L  S  R  P  N  A

CAGGAGCTTCCTAGCATGTACCAGCGCCTAGGGCTGGACTACGAGGAACGAGTGTTGCCGTCCATTGTCAACGAGGTGCTCAAGAGTGTG
GTCCTCGAAGGATCGTACATGGTCGCGGATCCCGACCTGATGCTCCTTGCTCACAACGGCAGGTAACAGTTGCTCCACGAGTTCTCACAC                 540

 Q  E  L  P  S  M  Y  Q  R  L  G  L  D  Y  E  E  R  V  L  P  S  I  V  N  E  V  L  K  S  V

GTGGCCAAGTTCAATGCCTCACAGCTGATCACCCAGCGGGCCCAGGTATCCCTGTTGATCCGCCGGGAGCTGACAGAGAGGGCCAAGGAC
CACCGGTTCAAGTTACGGAGTGTCGACTAGTGGGTCGCCCGGGTCCATAGGGACAACTAGGCGGCCCTCGACTGTCTCTCCCGGTTCCTG                 630

 V  A  K  F  N  A  S  Q  L  I  T  Q  R  A  Q  V  S  L  L  I  R  R  E  L  T  E  R  A  K  D

TTCAGCCTCATCCTGGATGATGTGGCCATCACAGAGCTGAGCTTTAGCCGAGAGTACACAGCTGCTGTAGAAGCCAAACAAGTGGCCCAG
AAGTCGGAGTAGGACCTACTACACCGGTAGTGTCTCGACTCGAAATCGGCTCTCATGTGTCGACGACATCTTCGGTTTGTTCACCGGGTC                 720

 F  S  L  I  L  D  D  V  A  I  T  E  L  S  F  S  R  E  Y  T  A  A  V  E  A  K  Q  V  A  Q

CAGGAGGCCCAGCGGGCCCAATTCTTGGTAGAAAAAGCAAAGCAGGAACAGCGGCAGAAAATTGTGCAGGCCGAGGGTGAGGCCGAGGCT
GTCCTCCGGGTCGCCCGGGTTAAGAACCATCTTTTTCGTTTCGTCCTTGTCGCCGTCTTTTAACACGTCCGGCTCCCACTCCGGCTCCGA                 810

 Q  E  A  Q  R  A  Q  F  L  V  E  K  A  K  Q  E  Q  R  Q  K  I  V  Q  A  E  G  E  A  E  A

GCCAAGATGCTTGGAGAAGCACTGAGCAAGAACCCTGGCTACATCAAACTTCGCAAGATTCGAGCAGCCCAGAATATCTCCAAGACGATC
CGGTTCTACGAACCTCTTCGTGACTCGTTCTTGGGACCGATGTAGTTTGAAGCGTTCTAAGCTCGTCGGGTCTTATAGAGGTTCTGCTAG                 900

 A  K  M  L  G  E  A  L  S  K  N  P  G  Y  I  K  L  R  K  I  R  A  A  Q  N  I  S  K  T  I
```

HMWG546

GCCACATCACAGAATCGTATCTATCTCACAGCTGACAACCTTGTGCTGAACCTACAGGATGAAAGTTTCACCAGGGGAAGTGACAGCCTC

CGGTGTAGTGTCTTAGCATAGATAGAGTGTCGACTGTTGGAACACGACTTGGATGTCCTACTTTCAAAGTGGTCCCCTTCACTGTCGGAG 990

A T S Q N R I Y L T A D N L V L N L Q D E S F T R G S D S L

ATCAAGGGTAAGAAATGAGCCTAGTCACCAAGAACTCCACCCCCACAAGAAGTGGATCTGCTTCTCCAGTTTTTGA 1066 (SEQ ID NO:16

TAGTTCCCATTCTTTACTCGGATCAGTGGTTCTTGAGGTGGGGGTGTTCTTCACCTAGACGAAGAGGTCAAAAACT

I K G K K . (SEQ ID NO:31)

 hNF. wo6

```
GGCACGAGATGACATCACTAAGTGGCCGATCTGCACAGAGCAGGCCAGGAGCAACCACACAGGCTTCCTGCACATGGACTGCGAGATCAA
---------------------------------------------------------------------------------------------+ 90
CCGTGCTCTACTGTAGTGATTCACCGGCTAGACGTGTCTCGTCCGGTCCTCGTTGGTGTGTCCGAAGGACGTGTACCTGACGCTCTAGTT

                                                                         M  D  C  E  I  K

GGGCCGCCCCTGCTGCATCGGCACCAAGGGCAGCTGTGAGATCACCACCCGGGAATACTGTGAGTTCATGCACGGCTATTTCCATGAGGA
---------------------------------------------------------------------------------------------+ 180
CCCGGCGGGGACGACGTAGCCGTGGTTCCCGTCGACACTCTAGTGGTGGGCCCTTATGACACTCAAGTACGTGCCGATAAAGGTACTCCT

 G  R  P  C  C  I  G  T  K  G  S  C  E  I  T  T  R  E  Y  C  E  F  M  H  G  Y  F  H  E  E

AGCAACACTCTGCTCCCAGGTGAGGCGAGGCAGGCCTGGAGTAGTGGAGGAGAGGACGCTGGGCATGGCAGCCTGCTGGGGCCGGGGCTC
---------------------------------------------------------------------------------------------+ 270
TCGTTGTGAGACGAGGGTCCACTCCGCTCCGTCCGGACCTCATCACCTCCTCTCCTGCGACCCGTACCGTCGGACGACCCCGGCCCCGAG

 A  T  L  C  S  Q  V  R  R  G  R  P  G  V  V  E  E  R  T  L  G  M  A  A  C  W  G  R  G  S

ACGCACTCCCTCCCATGTCGGAGCCTCAGACTCAGGCTGCTTCTGGGGCGCTGAGCACCATATGCCCATTCCCAGGTGCACTGTTTTGGA
---------------------------------------------------------------------------------------------+ 360
TGCGTGAGGGAGGGTACAGCCTCGGAGTCTGAGTCCGACGAAGACCCCGCGACTCGTGGTATACGGGTAAGGGTCCACGTGACAAAACCT

 R  T  P  S  H  V  G  A  S  D  S  G  C  F  W  G  A  E  H  H  M  P  I  P  R  C  T  V  L  D

CAAGGTGTGTTGGGCTGCTGCCTTCCTCAACCCTGAGGTCCCAGATCAGTTTTACAGGTCTGGCTGTCTCTTTTCCTACATGTTGGGTAA
---------------------------------------------------------------------------------------------+ 450
GTTCCACACAACCCGACGACGGAAGGAGTTGGGACTCCAGGGTCTAGTCAAAATGTCCAGACCGACAGAGAAAAGGATGTACAACCCATT

 K  V  C  W  A  A  A  F  L  N  P  E  V  P  D  Q  F  Y  R  S  G  C  L  F  S  Y  M  L  G  K

GAGGTCCTCAATGCCCCCGAACCCGACCCCTGTGATGGACACCCAGGCGGACCCCTGGGGAAAGGTTCCTGGGCCAGGGTATGGTCGGTC
---------------------------------------------------------------------------------------------+ 540
CTCCAGGAGTTACGGGGGCTTGGGCTGGGGACACTACCTGTGGGTCCGCCTGGGGACCCCTTTCCAAGGACCCGGTCCCATACCAGCCAG

 R  S  S  M  P  P  N  P  T  P  V  M  D  T  Q  A  D  P  W  G  K  V  P  G  P  G  Y  G  R  S

CAACCTGCCGAAGACTACTGCTCCTGAAGTGTCTGGATGAAGGCCGCTGCCTGGTGTGTCCCTCCCCCAGTGTGGGTGCACTGCCCTCGG
---------------------------------------------------------------------------------------------+ 630
GTTGGACGGCTTCTGATGACGAGGACTTCACAGACCTACTTCCGGCGACGGACCACACAGGGAGGGGGTCACACCCACGTGACGGGAGCC

 N  L  P  K  T  T  A  P  E  V  S  G  .  (SEQ ID NO:32)

TGTCTTGTGGGTCTTTTCAAATGACATCCCTGAAGGGGACCTGGAGGAAGGTGGTCCGGCTGGCACCCTATCGC     (SEQ ID NO:17)
------------------------------------------------------------------------------> 704
ACAGAACACCCAGAAAAGTTTACTGTAGGGACTTCCCCTGGACCTCCTTCCACCAGGCCGACCGTGGGATAGCG
```

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 13, ~~line~~ Table 1

B. IDENTIFICATION OF DEPOSIT — Further deposits are identified on an additional sheet ☐

Name of depositary institution
American Type Culture Collection

Address of depositary institution *(including postal code and country)*
12301 Parklawn Drive
Rockville, Maryland 20852
United States of America

Date of deposit	Accession Number
May 16, 1997	209053

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* — This information is continued on an additional sheet ☐

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only

☑ This sheet was received with the international application

Authorized officer Yvette Simms
Paralegal Specialist
IAPD-PCT Operations
(703) 305-9748

For International Bureau use only

☐ This sheet was received by the International Bureau on:

Authorized officer

Form PCT/RO/134 (July 1992)

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 13, line Table 1

B. IDENTIFICATION OF DEPOSIT — Further deposits are identified on an additional sheet ☐

Name of depositary institution

American Type Culture Collection

Address of depositary institution *(including postal code and country)*

12301 Parklawn Drive
Rockville, Maryland 20852
United States of America

Date of deposit	Accession Number
May 16, 1997	209053

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* — This information is continued on an additional sheet ☐

EUROPE

In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused of withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only

☑ This sheet was received with the international application

Authorized officer Yvette Simms
Paralegal Specialist
PCT-PCT Operations

For International Bureau use only

☐ This sheet was received by the International Bureau on:

Authorized officer

Form PCT/RO/134 (July 1992)

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 13 line/ Table 1

B. IDENTIFICATION OF DEPOSIT — Further deposits are identified on an additional sheet ☐

Name of depositary institution
American Type Culture Collection

Address of depositary institution *(including postal code and country)*
12301 Parklawn Drive
Rockville, Maryland 20852
United States of America

Date of deposit	Accession Number
May 16, 1997	209054

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* — This information is continued on an additional sheet ☐

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☑ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer Yvette Simms Paralegal Specialist APD-PCT Operations	Authorized officer

Form PCT/RO/134 (July 1992)

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 13, ~~line~~ Table 1.

B. IDENTIFICATION OF DEPOSIT — Further deposits are identified on an additional sheet ☐

Name of depositary institution

American Type Culture Collection

Address of depositary institution *(including postal code and country)*

12301 Parklawn Drive
Rockville, Maryland 20852
United States of America

Date of deposit	Accession Number
May 16, 1997	**209054**

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* — This information is continued on an additional sheet ☐

EUROPE
In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only

☑ This sheet was received with the international application

Authorized officer

Yvette Simms
Paralegal Specialist
IAPD-PCT Operations
(703) 305-3746

For International Bureau use only

☐ This sheet was received by the International Bureau on:

Authorized officer

What is claimed is:

1. An isolated polynucleotide comprising a nucleotide sequence that has at least 80% identity over its entire length to a nucleotide sequence encoding the receptor polypeptide of SEQ ID NO:Y; or a nucleotide sequence complementary to said isolated polynucleotide.

2. The polynucleotide of claim 1 wherein said polynucleotide comprises the nucleotide sequence contained in SEQ ID NO:X encoding the receptor polypeptide of SEQ ID NO:Y.

3. The polynucleotide of claim 1 wherein said polynucleotide comprises a nucleotide sequence that is at least 80% identical to that of SEQ ID NO:X over its entire length.

4. The polynucleotide of claim 3 which is polynucleotide of SEQ ID NO: X.

5. The polynucleotide of claim 1 which is DNA or RNA.

6. A DNA or RNA molecule comprising an expression system, wherein said expression system is capable of producing a receptor polypeptide comprising an amino acid sequence, which has at least 80% identity with the polypeptide of SEQ ID NO:Y when said expression system is present in a compatible host cell.

7. A host cell comprising the expression system of claim 6.

8. A process for producing a receptor polypeptide comprising culturing a host of claim 7 under conditions sufficient for the production of said polypeptide and recovering the polypeptide from the culture.

9. A process for producing a cell which produces a receptor polypeptide thereof comprising transforming or transfecting a host cell with the expression system of claim 6 such that the host cell, under appropriate culture conditions, produces a receptor polypeptide.

10. A receptor polypeptide comprising an amino acid sequence which is at least 80% identical to the amino acid sequence of SEQ ID NO:Y over its entire length.

11. The polypeptide of claim 10 which comprises the amino acid sequence of SEQ ID NO:Y.

12. An antibody immunospecific for the receptor polypeptide of claim 10.

13. A method for the treatment of a subject in need of enhanced activity or expression of receptor polypeptide of claim 10 comprising:

(a) administering to the subject a therapeutically effective amount of an agonist to said receptor; and/or

(b) providing to the subject polynucleotide of claim 1 in a form so as to effect production of said receptor activity in vivo.

14. A method for the treatment of a subject having need to inhibit activity or expression of the receptor polypeptide of claim 10 comprising:

(a) administering to the subject a therapeutically effective amount of an antagonist to said receptor; and/or

(b) administering to the subject a nucleic acid molecule that inhibits the expression of the nucleotide sequence encoding said receptor; and/or

(c) administering to the subject a therapeutically effective amount of a polypeptide that competes with said receptor for its ligand.

15. A process for diagnosing a disease or a susceptibility to a disease in a subject related to expression or activity of the receptor polypeptide of claim 10 in a subject comprising:

(a) determining the presence or absence of a mutation in the nucleotide sequence encoding said receptor polypeptide in the genome of said subject; and/or

(b) analyzing for the presence or amount of the receptor polypeptide expression in a sample derived from said subject.

16. A method for identifying agonists to the receptor polypeptide of claim 10 comprising:

(a) contacting cells produced by claim 9 with a candidate compound; and

(b) determining whether the candidate compound effects a signal generated by activation of the receptor polypeptide.

17. An agonist identified by the method of claim 16.

18. The method for identifying antagonists to the receptor polypeptide of claim 10 comprising:

(a) contacting said cell produced by claim 9 with an agonist; and

(b) determining whether the signal generated by said agonist is diminished in the presence of a candidate compound.

19. An antagonist identified by the method of claim 18.

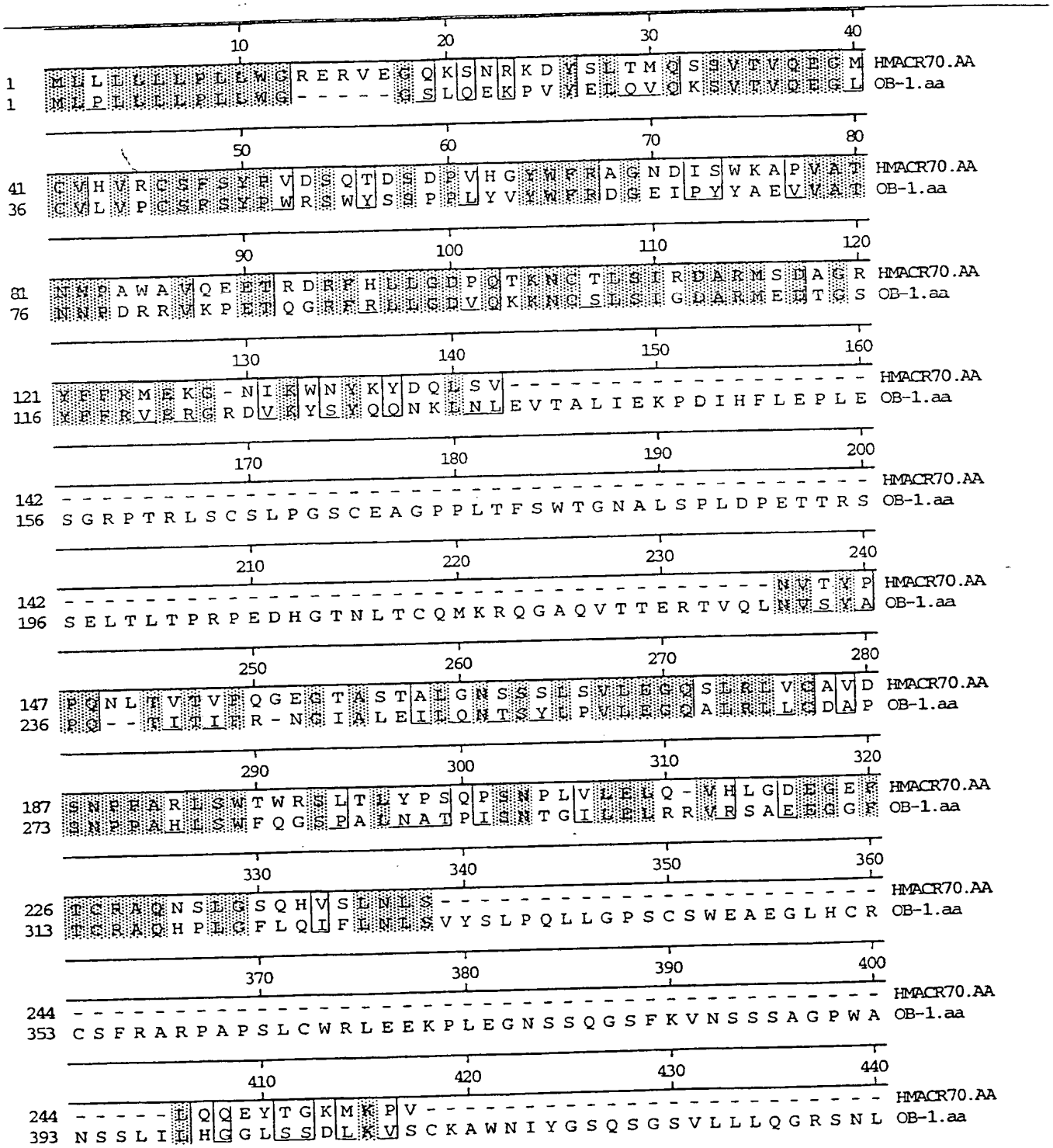
20. An isolated receptor polynucleotide comprising a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence having at least 80% identity to a nucleotide sequence encoding the receptor polypeptide expressed by the cDNA insert deposited at the ATCC; and

(b) a nucleotide sequence complementary to the nucleotide sequence of (a).

21. A recombinant host cell produced by a method of Claim 9 or a membrane thereof expressing a receptor polypeptide.

FIGURE 1


10 20 30 40
1 MLLLLLLPLLWGRERVEGQKSNRKDYSLTMQSSVTVQEGML HMACR70.AA
1 MLPLLLLPLLWG-----GSLQEKPVYELQVQKSVTVQEGLL OB-1.aa
50 60 70 80
41 CVHVRCSFSYFVDSQTDSDPVHGYWFRAGNDISWKAPVATL HMACR70.AA
36 CVLVPCSRSYPWRSWYSSPPLYVYWFRDGEIPYYAEVVATL OB-1.aa
90 100 110 120
81 NMPAWAVQEETRDRFHLLGDPQTKNCTLSIRDARMSDAGRL HMACR70.AA
76 NNRDRRVKPETQGRFRLLGDVQKKNCSLSIGDARMEDTGSL OB-1.aa
130 140 150 160
121 YEFRMEKG-NIKWNYKYDQLSV---------------- HMACR70.AA
116 YFFRVERGRDVKYSYQQNKLNLEVTALIEKPDIHFLEPLE OB-1.aa
170 180 190 200
142 ---------- HMACR70.AA
156 SGRPTRLSCSLPGSCEAGPPLTFSWTGNALSPLDPETTRS OB-1.aa
210 220 230 240
142 ----------NVTYP HMACR70.AA
196 SELTLTPRPEDHGTNLTCQMKRQGAQVTTERTVQLNVSYA OB-1.aa
250 260 270 280
147 PQNLTVTVEQGETASTALGNSSSLSVLEGQSLRLVCAVD HMACR70.AA
236 PQ--TITIER-NGIALEILQNTSYLPVLEGQALRLLCDAP OB-1.aa
290 300 310 320
187 GNPEARLSWTWRSLTLYPSQPSNPLVLELQ-VHLGDEGEF HMACR70.AA
273 GNPPAHLSWFQGSPALNATPISNTGILELRRVRSAEDGGFY OB-1.aa
330 340 350 360
226 TCRAQNSLGSQHVSLNLS---------- HMACR70.AA
313 TCRAQHPLGFLQIFHNLSVYSLPQLLGPSCSWEAEGLHCR OB-1.aa
370 380 390 400
244 ---------- HMACR70.AA
353 CSFRARPAPSLCWRLEEKPLEGNSSQGSFKVNSSSAGPWA OB-1.aa
410 420 430 440
244 -----LQQEYTGKMKPV---------- HMACR70.AA
393 NSSLILHGGLSSDLKVSCKAWNIYGSQSGSVLLLQGRSNL OB-1.aa

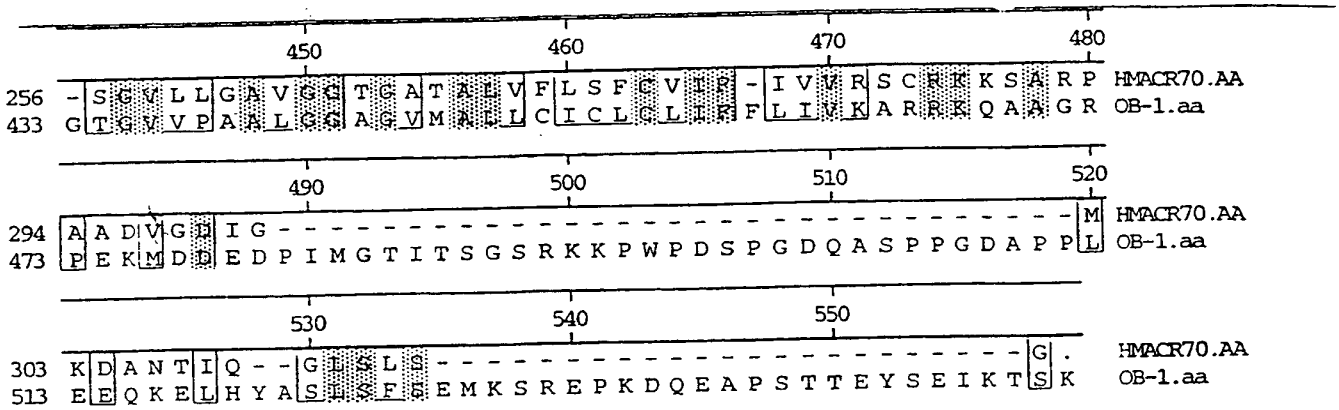


FIGURE 2

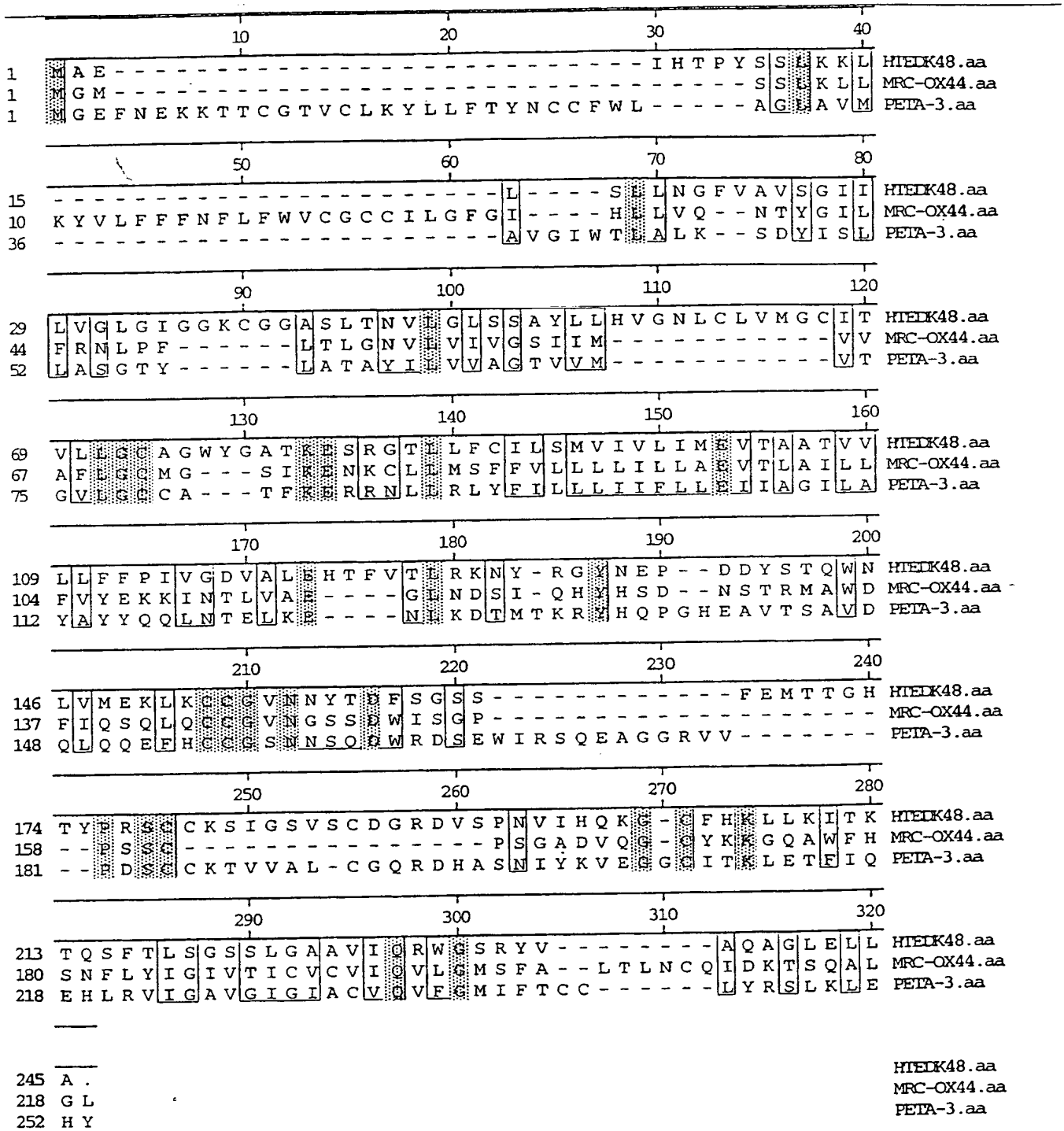


FIGURE 3

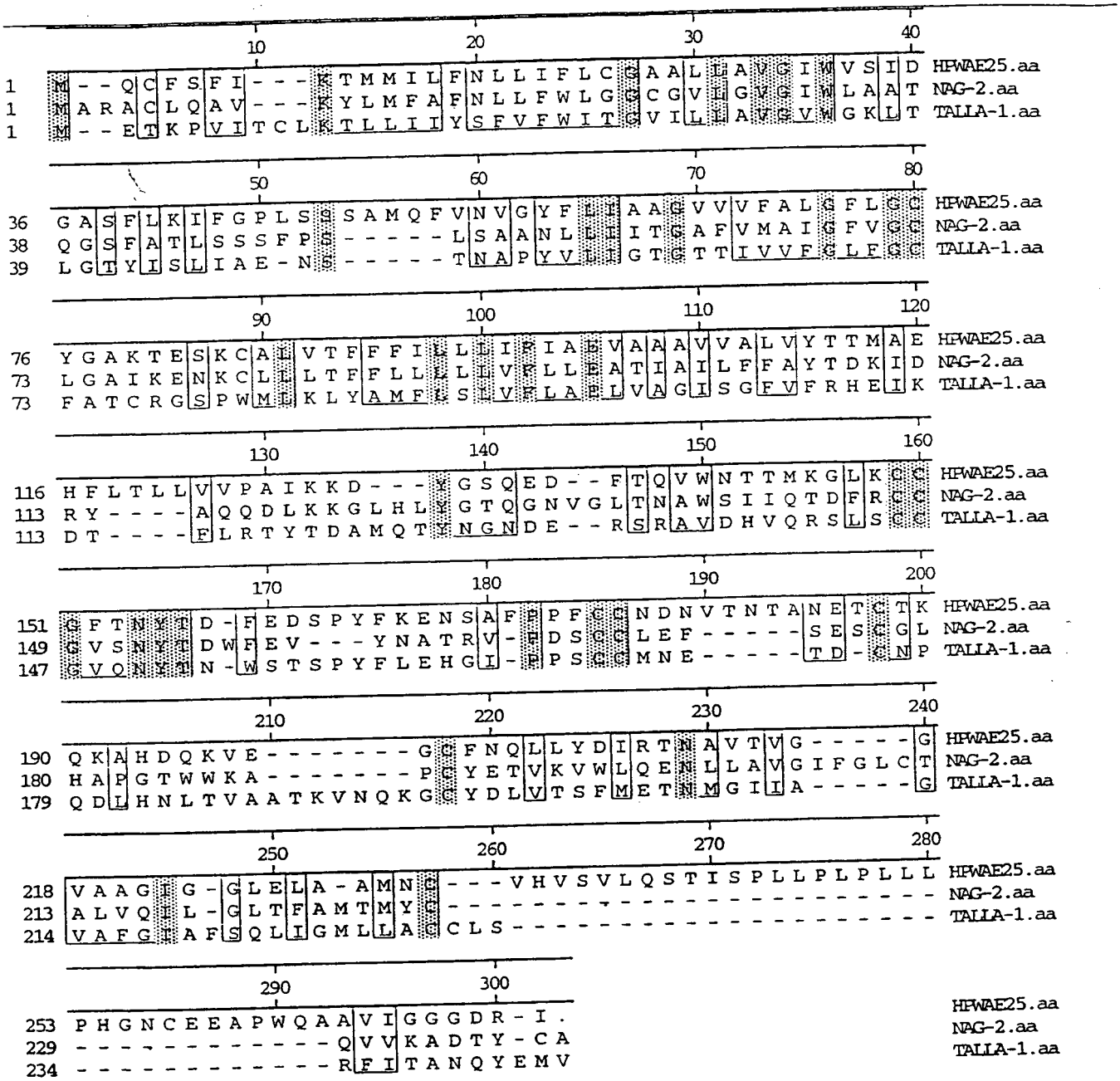


FIGURE 4

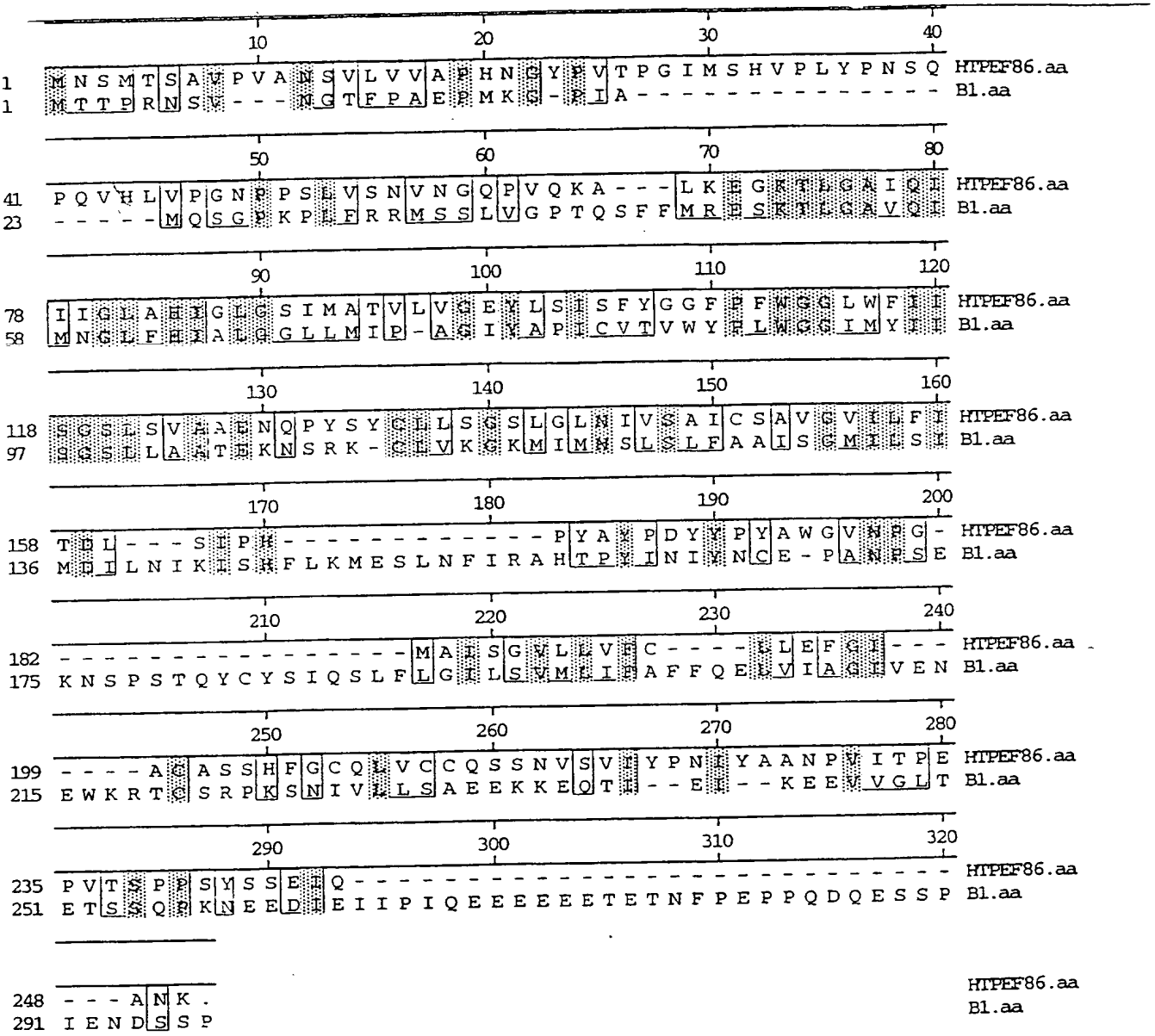

10 20 30 40
1 MNSMTSAVPVANSVLVVAPHNGYPVTPGIMSHVPLYPNSQ HTPEF86.aa
1 MTTPRNSV---NGTFPAEPMKG-PIA---------------- B1.aa
50 60 70 80
41 PQVHLVPGNPPSLVSNVNGQPVQKA---LKEGKTLGAIQI HTPEF86.aa
23 -----MQSGPKPLFRRMSSLVGPTQSFFMREGKTLGAVQI B1.aa
90 100 110 120
78 IIGLAHIGLGSIMATVLVGEYLSISFYGGFPFWGGLWFIII HTPEF86.aa
58 MNGLFHIALGGLLMIP-AGIVAPICVTVWYHLWGGIMYIII B1.aa
130 140 150 160
118 SGSLSVAAENQPYSYGLLSGSLGLNIVSAICSAVGVILFI HTPEF86.aa
97 SGSLLAATEKNSRK-GLVKGKMIMMSLSLFAAISGMILSI B1.aa
170 180 190 200
158 TDL---SIPH----------PYAYPDYYPYAWGVNPG- HTPEF86.aa
136 MDILNIKISHFLKMESLNFIRAHTPYINIYNCE-PANPSE B1.aa
210 220 230 240
182 ----------MAISGVLLVFC----LLEFGI--- HTPEF86.aa
175 KNSPSTQYCYSIQSLFLGILSVMLIPAFFQELVIAGIVEN B1.aa
250 260 270 280
199 ----AGASSHFGCQLVCCQSSNVSVIYPNIYAANPVITPE HTPEF86.aa
215 EWKRTGSRPKSNIVLLSAEEKKEQTI--EL--KEEVVGLT B1.aa
290 300 310 320
235 PVTSPPSYSSEIQ---------------------------- HTPEF86.aa
251 ETSSQPKNEEDIEIIPIQEEEEEETETNFPEPPQDQESSP B1.aa
248 ---ANK. HTPEF86.aa
291 IENDSSP B1.aa

FIGURE 5

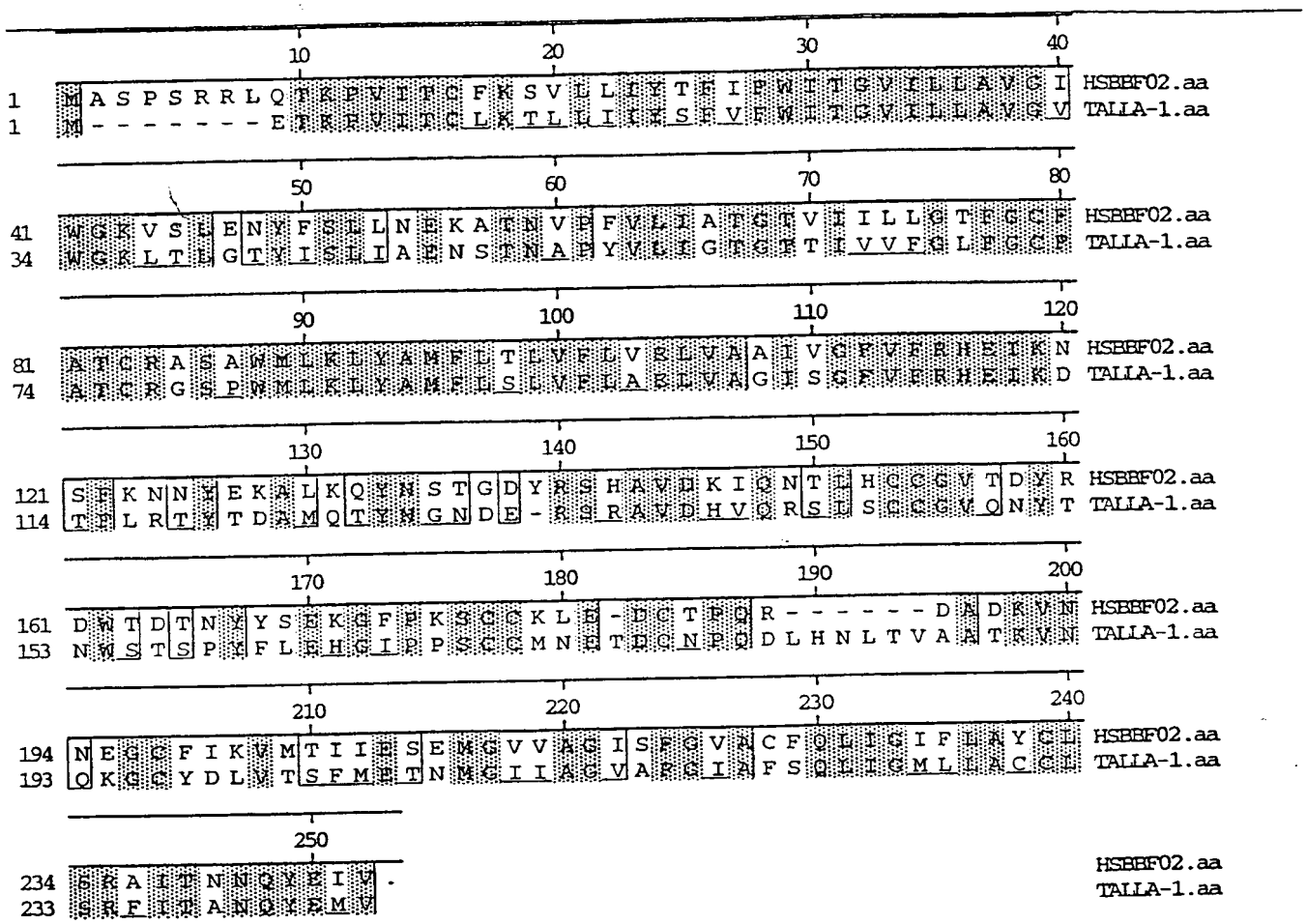


FIGURE 6

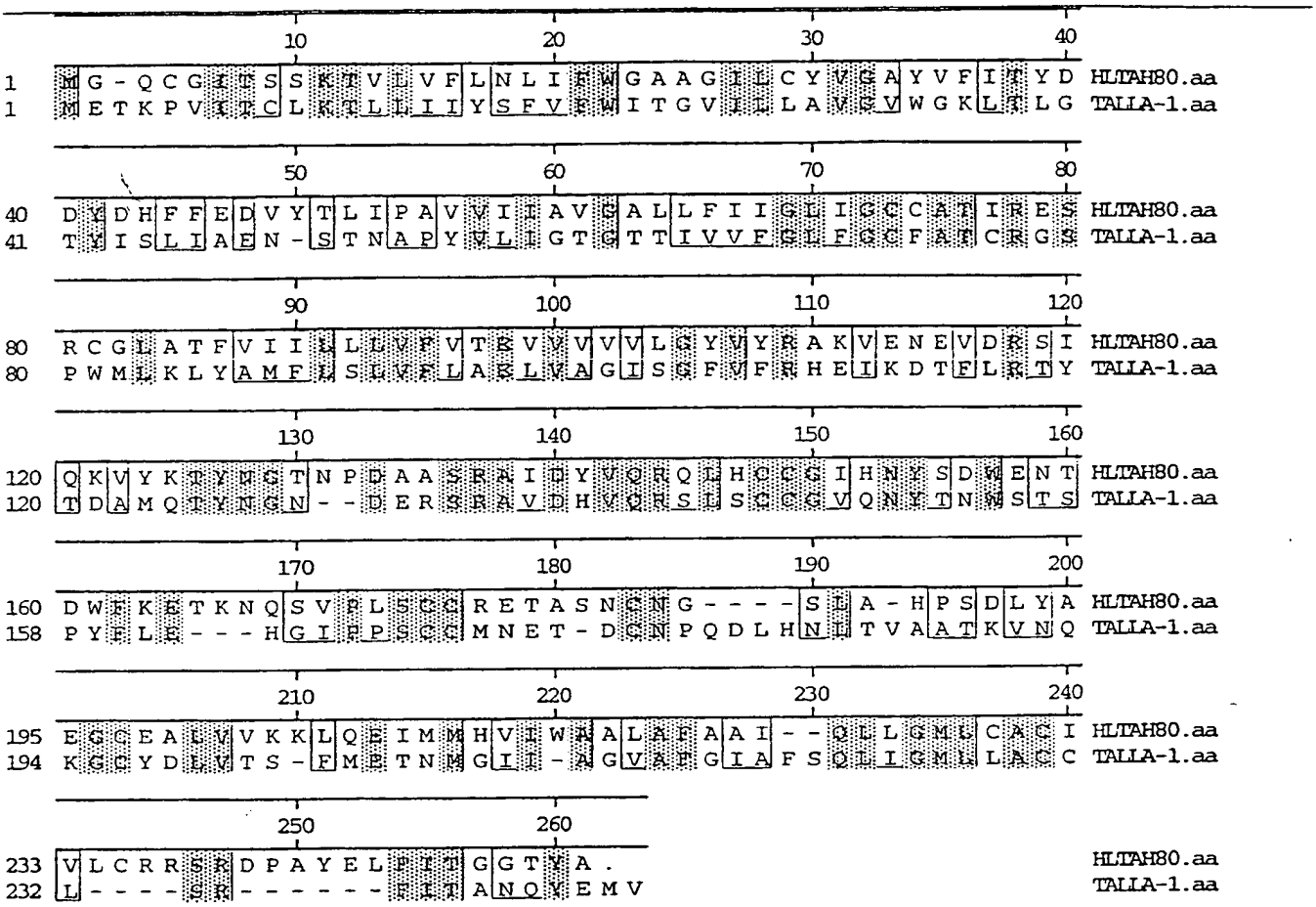


FIGURE 7

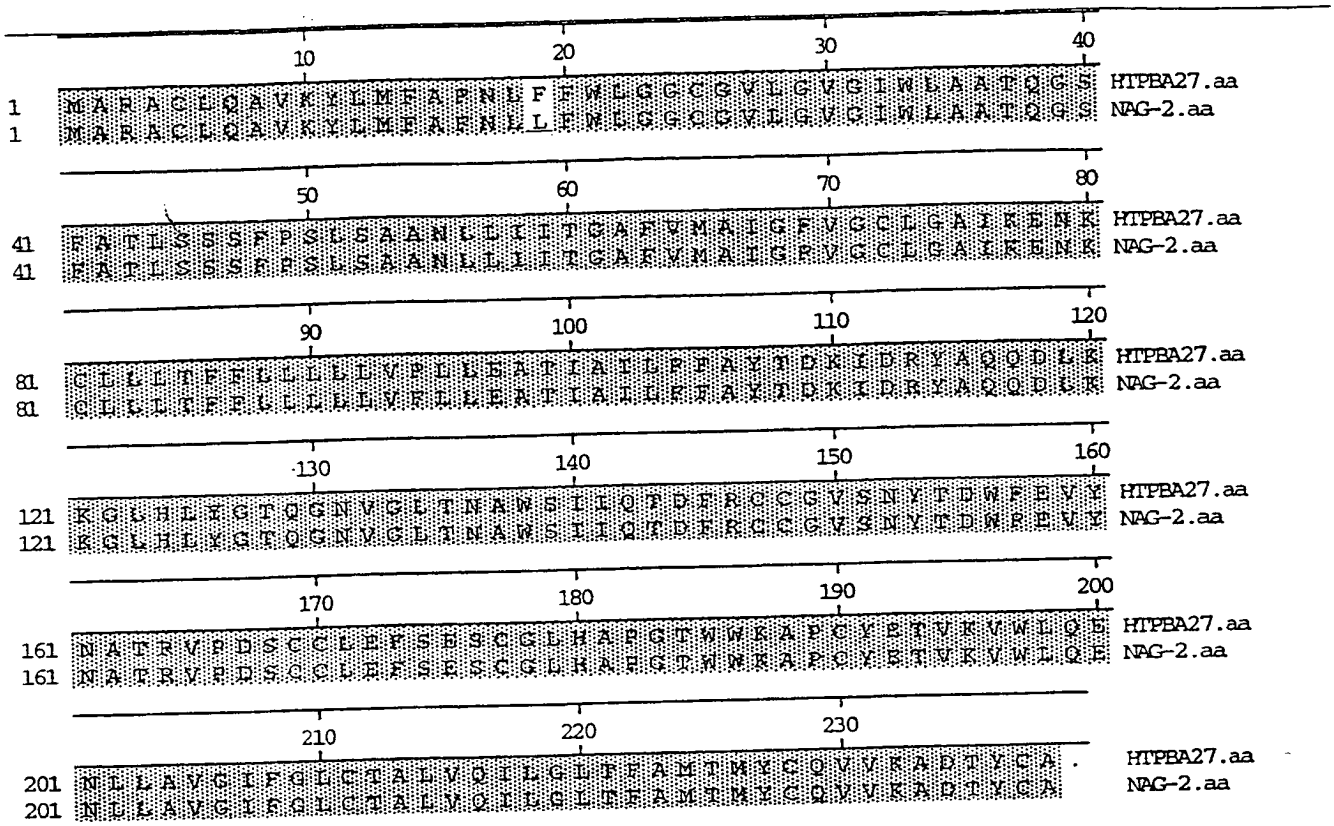

10
20
30
40
HTPBA27.aa
NAG-2.aa
50
60
70
80
HTPBA27.aa
NAG-2.aa
90
100
110
120
HTPBA27.aa
NAG-2.aa
130
140
150
160
HTPBA27.aa
NAG-2.aa
170
180
190
200
HTPBA27.aa
NAG-2.aa
210
220
230
HTPBA27.aa
NAG-2.aa

FIGURE 8

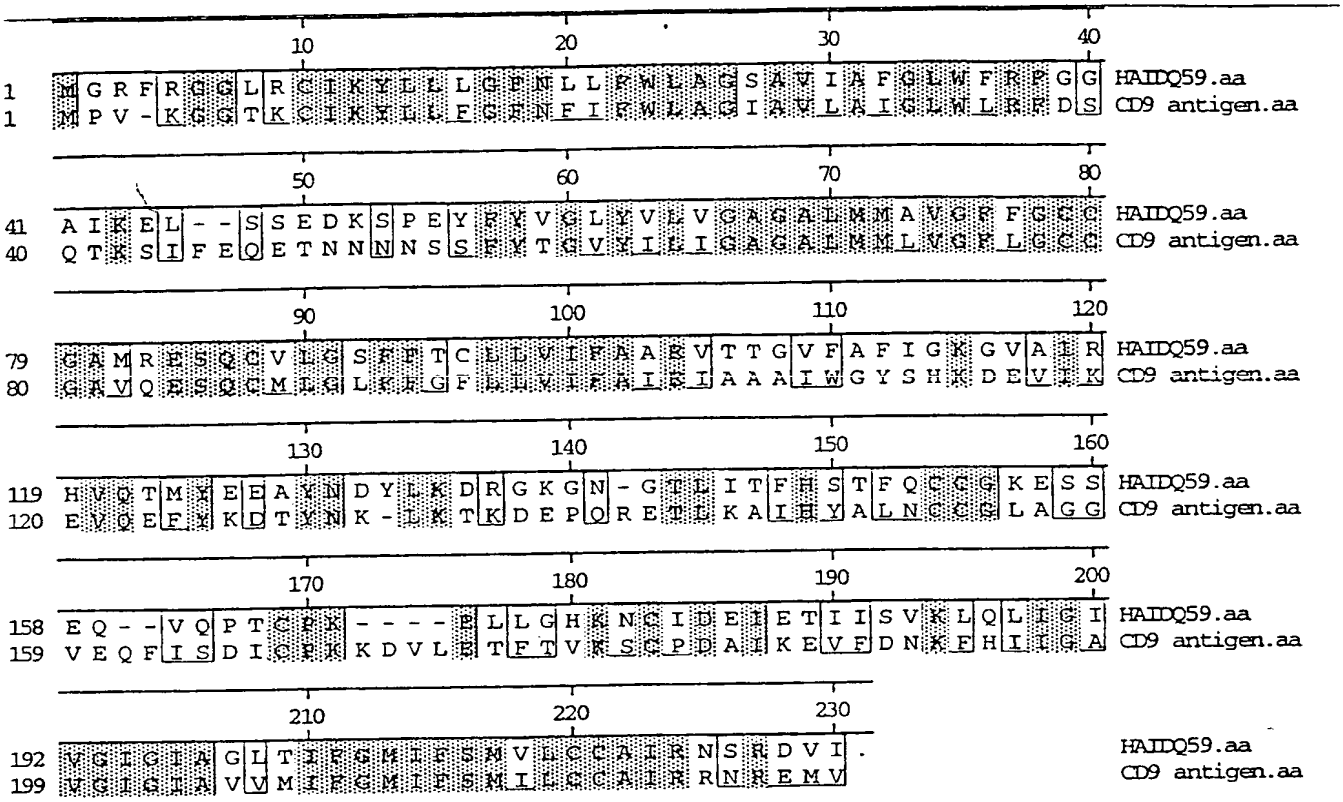


FIGURE 9

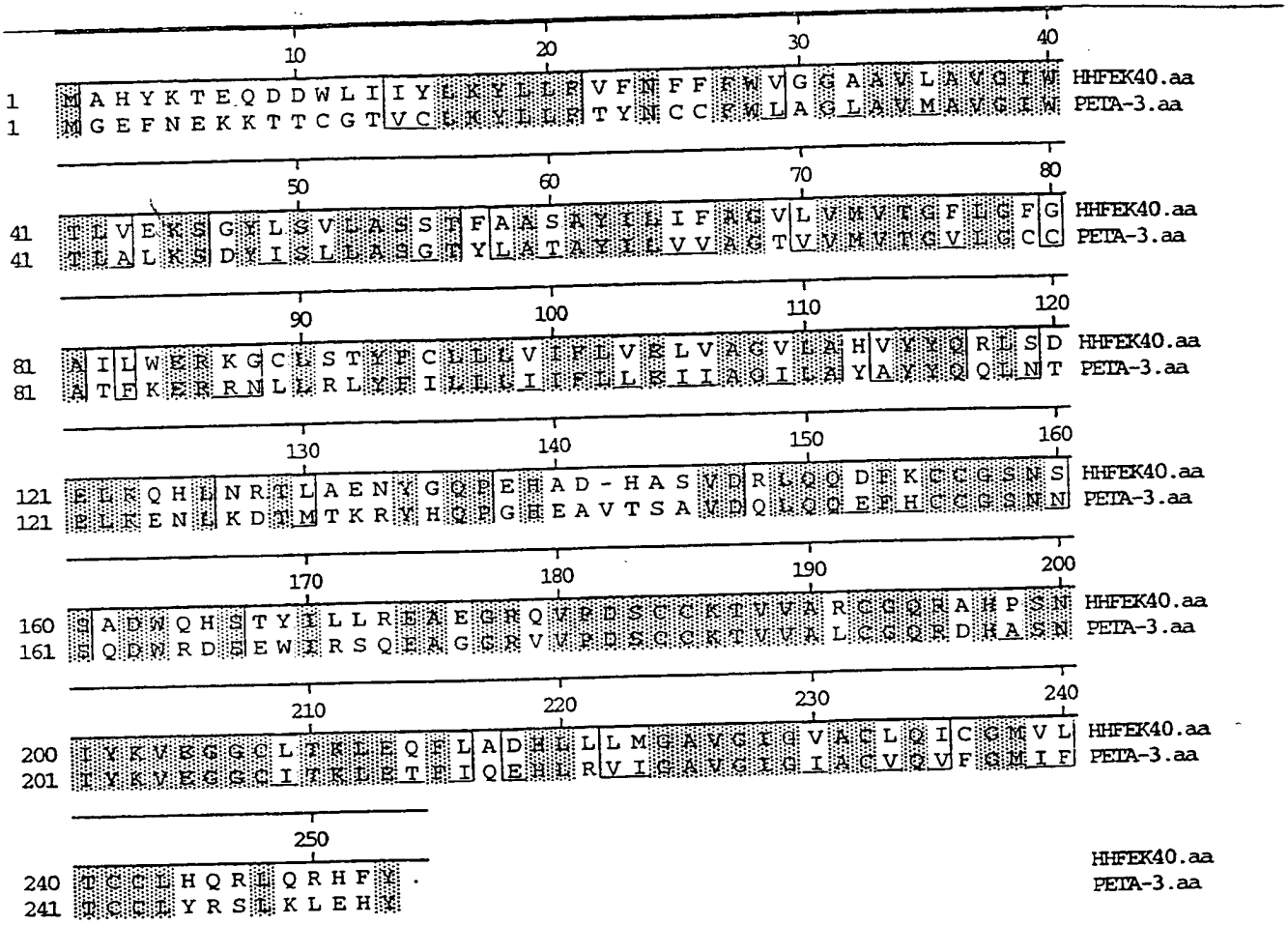

10 20 30 40
1 MAHYKTEQDDWLIIYLKYLLRVFNFFFWVGGAAVLAVGIW HHFEK40.aa
1 MGEFNEKKTTCGTVCLKYLLRTYNCCKWLAGLAVMAVGIW PETA-3.aa
50 60 70 80
41 TLVEKSGYLSVLASSTFAASAYLLIFAGVLVMVTGFLGFG HHFEK40.aa
41 TLALKSDYISLLASGTYLATAYYLLVVAGTVVMVTGVLGC PETA-3.aa
90 100 110 120
81 AILWERKGCLSTYFCLLLVIFLVRLVAGVLAHVYYQRLSD HHFEK40.aa
81 ATFKERRNLLRLYFILLLIIFLLRIIAQILAYAYYQQLNT PETA-3.aa
130 140 150 160
121 ELKQHLNRTLAENYGQPEHADHASVDRLQQDLKCCGSNS HHFEK40.aa
121 ELKENLKDTMTKRYHQPGHEAVTSAVDQLQQEPHCCGSNN PETA-3.aa
170 180 190 200
160 SADWQHSTYILLREAEGRQVPDSCCKTVVARCCQRAHPSN HHFEK40.aa
161 SQDWRDSEWIRSQEAGGRVVPDSCCKTVVALCGQRDHASN PETA-3.aa
210 220 230 240
200 LYKVEGGCLTKLEQFLADHLLLMGAVGLQVACLQICGMVL HHFEK40.aa
201 TYKVEGGCITKLETFIQEHLRVIGAVGLQIACVQVFGMIF PETA-3.aa
250
240 TCCLHQRLQRHFY. HHFEK40.aa
241 TCCIYRSLKLEHY PETA-3.aa

FIGURE 10

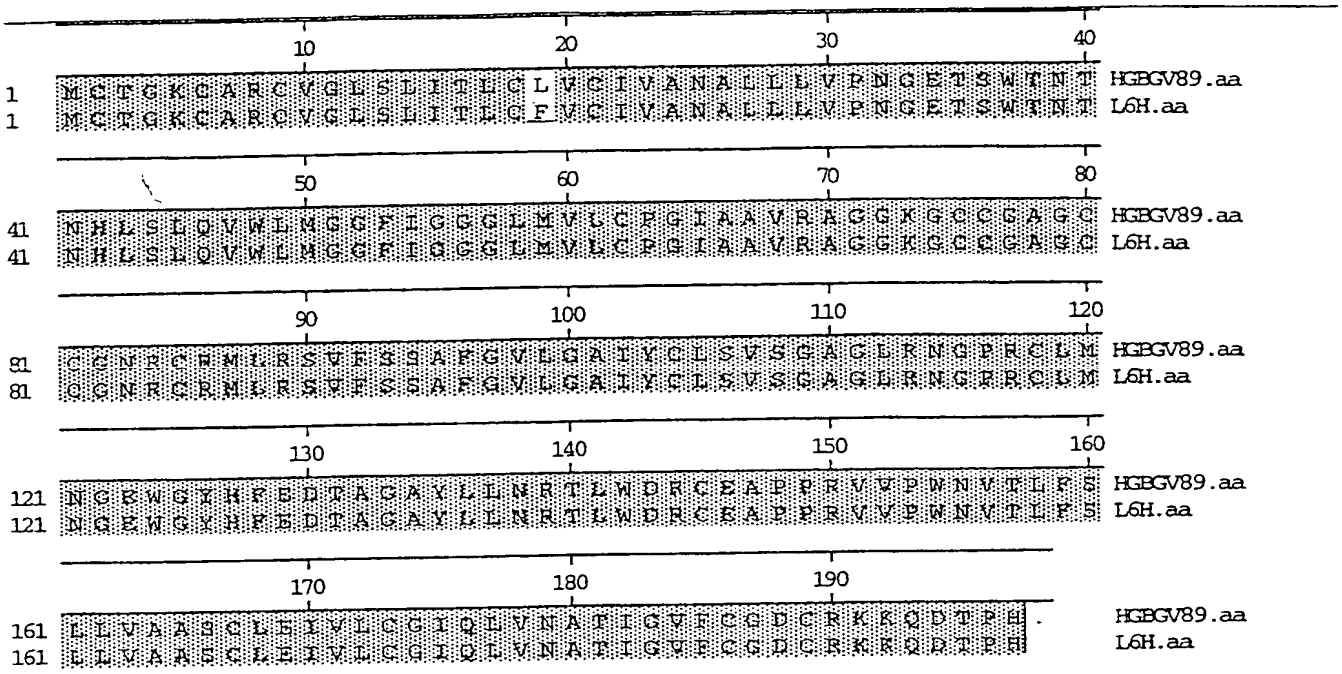


FIGURE 11

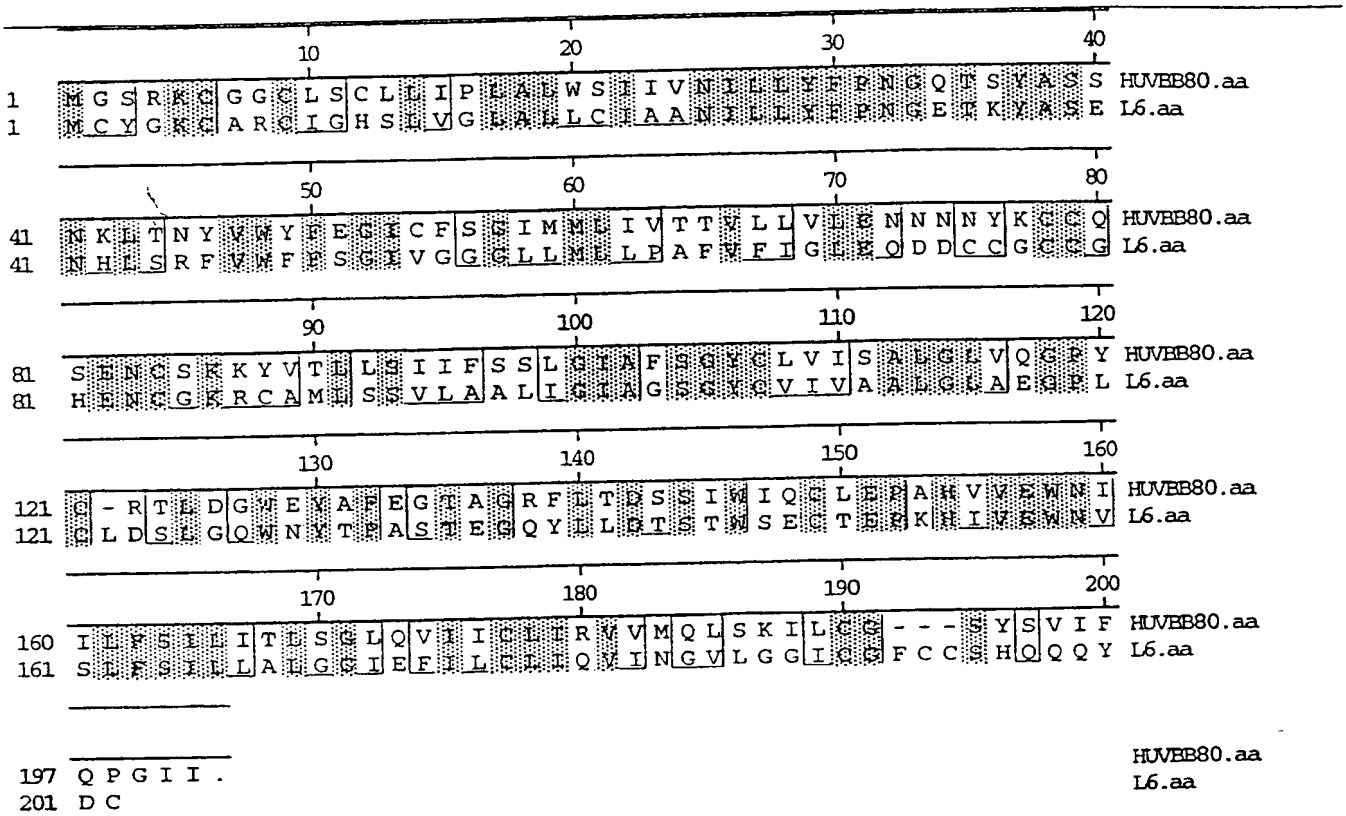

10 20 30 40
1 M G S R K C G G C L S C L L I P L A L W S I I V N I L L Y F R N G Q T S Y A S S HUVBB80.aa
1 M C Y G K C A R C I G H S L V G L A L L C I A A N I L L Y F R N G E T K Y A S E L6.aa
50 60 70 80
41 K K L T N Y V W Y F E G I C F S S I M M L I V T T Y L L V L E N N N N Y K G C Q HUVBB80.aa
41 N H L S R F V W F F S G I V G G L L M L L P A F W F I G L E Q D D C C G C C G L6.aa
90 100 110 120
81 S E N C S K K Y V T L L S I I F S S L G T A F S Q Y C L V I S A L G L V Q G R Y HUVBB80.aa
81 H E N C G K R C A M L S S V L A A L I G I A G S Q Y C V I V A A L G L A E G P L L6.aa
130 140 150 160
121 C - R T L D G W E Y A F E G T A G R F L T D S S I W I Q C L E P A K V Y E W N I HUVBB80.aa
121 C L D S L G Q W N Y T P A S T E G Q Y L L D T S T W S E C T E P K H I V E W N V L6.aa
170 180 190 200
160 I L F S I L I T L S G L Q V I I C L I R W V M Q L S K I L C G - - - S Y S V I F HUVBB80.aa
161 S L F S I L A L G G I E F I L C L I Q Y I N G V L G G I C G F C C S H Q Q Q Y L6.aa
197 Q P G I I . HUVBB80.aa
201 D C L6.aa

FIGURE 12

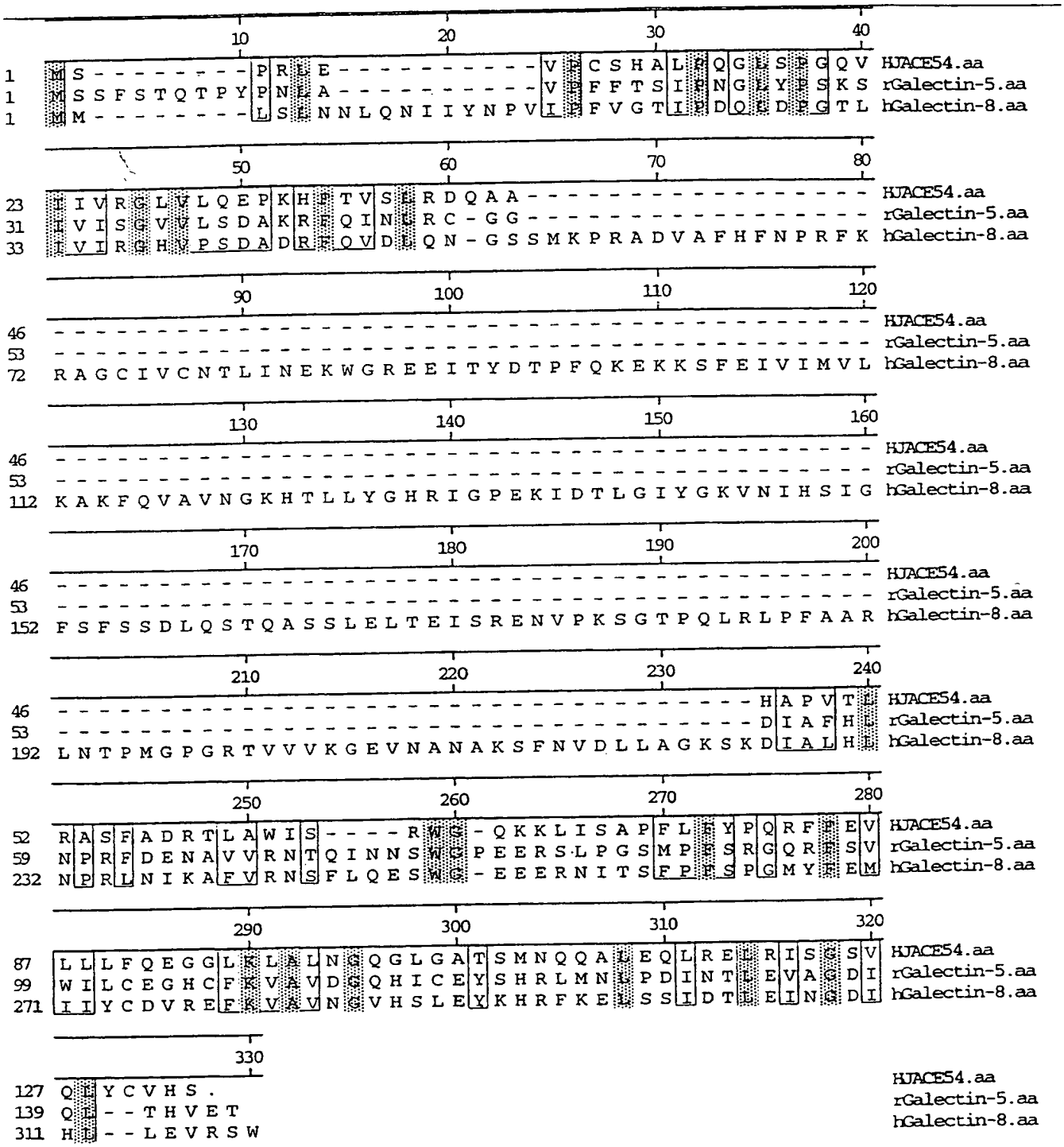


FIGURE 13

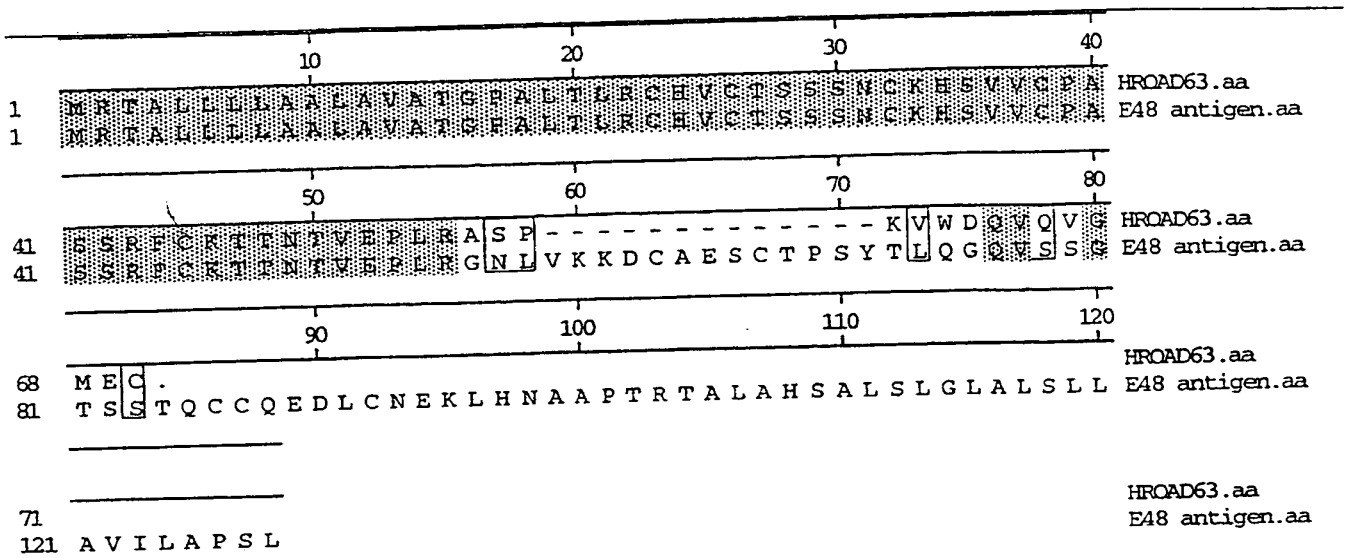


FIGURE 14

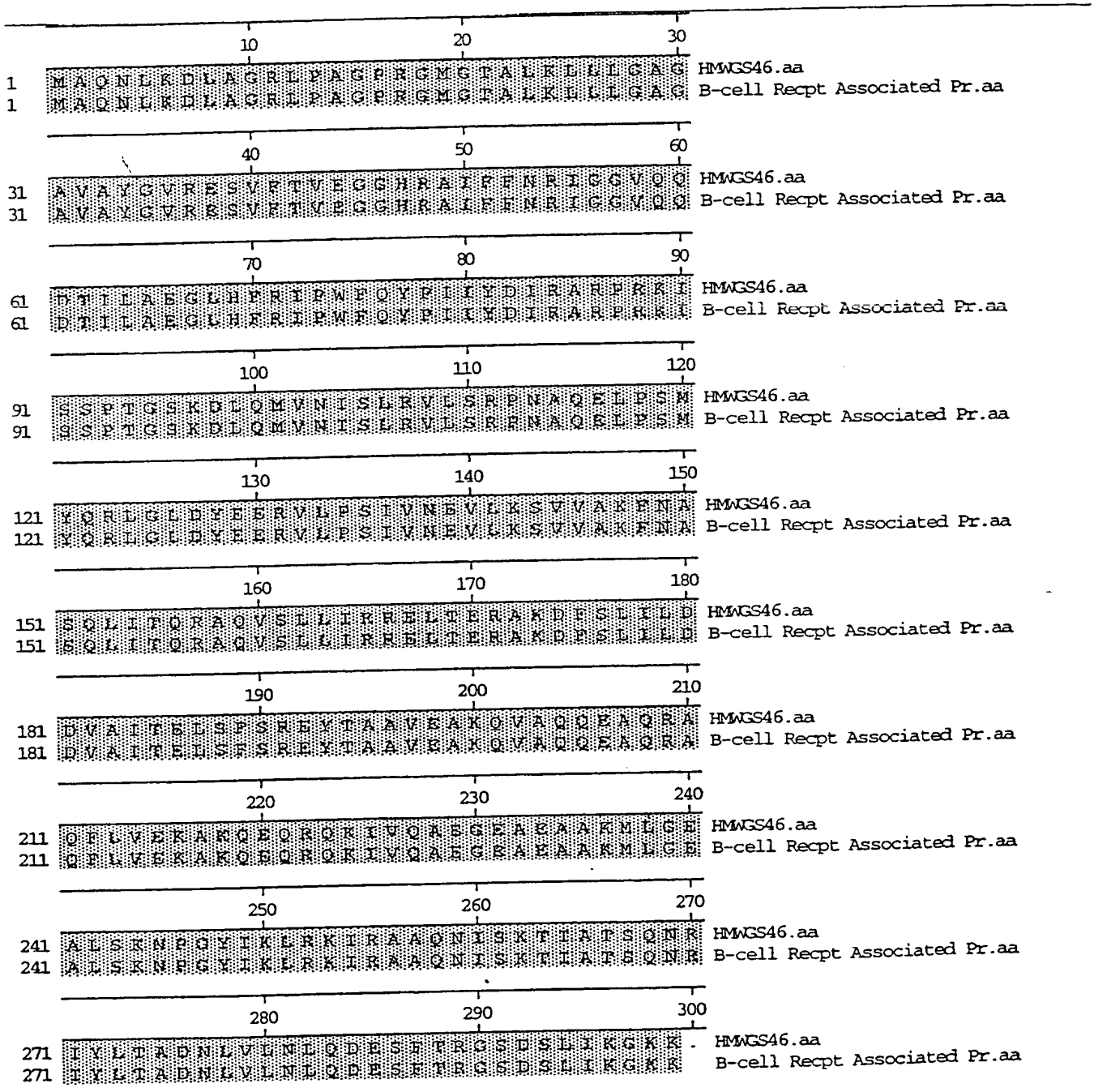
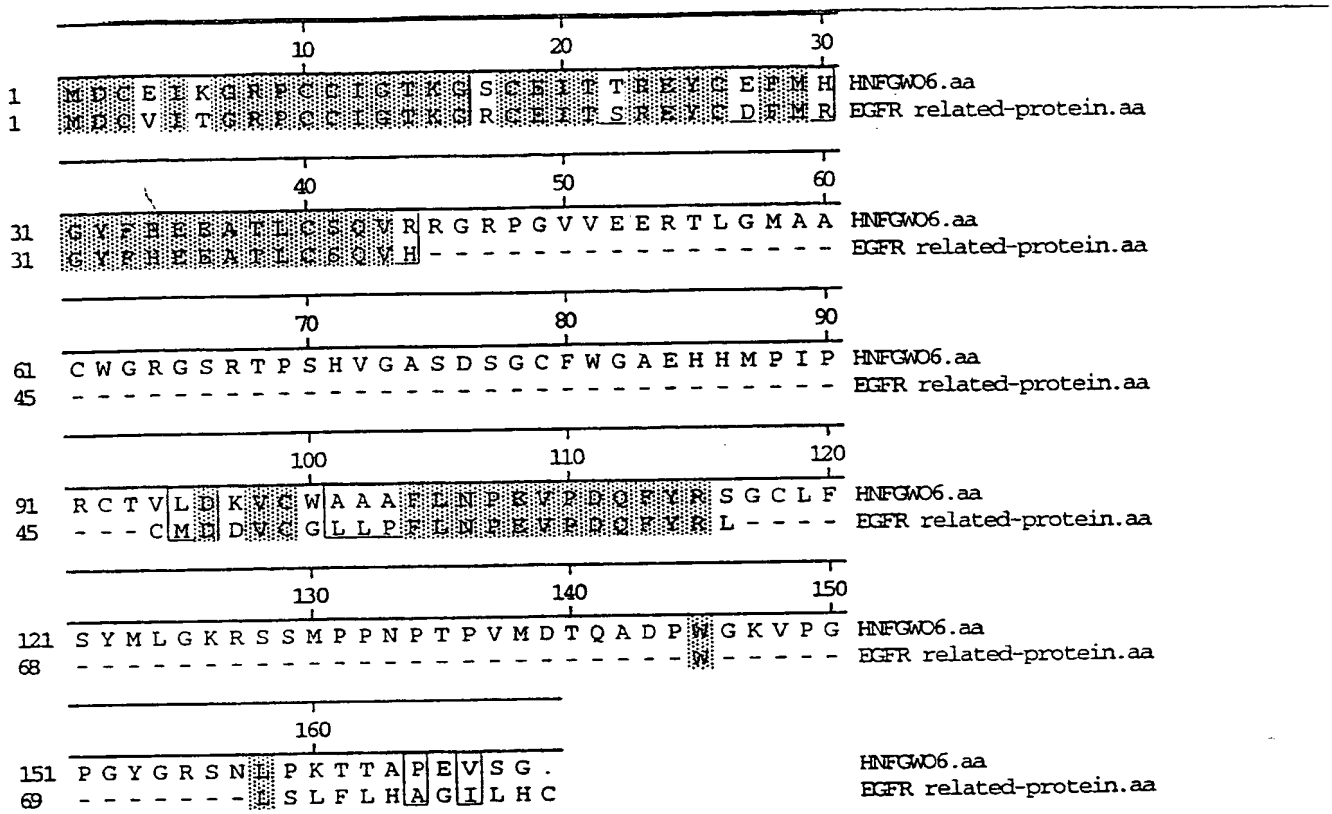


FIGURE 15



PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, 15/85, 5/10, C07K 14/705, 16/28, C12Q 1/68, A61K 38/17, 48/00

A3

(11) International Publication Number: WO 98/31799

(43) International Publication Date: 23 July 1998 (23.07.98)

(21) International Application Number: PCT/US98/00959

(22) International Filing Date: 21 January 1998 (21.01.98)

(30) Priority Data:

60/034,204	21 January 1997 (21.01.97)	US
60/034,205	21 January 1997 (21.01.97)	US

(71) Applicant *(for all designated States except US)*: HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: NI, Jian [CN/US]; 5502 Manorfield Road, Rockville, MD 20853 (US). GENTZ, Reiner, L. [DE/US]; 13404 Fairland Park Drive, Silver Spring, MD 20904 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US).

(74) Agents: BROOKES, A., Anders et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
5 November 1998 (05.11.98)

(54) Title: POLYNUCLEOTIDES AND POLYPEPTIDES ENCODING RECEPTORS

(57) Abstract

Receptor polypeptides and polynucleotides and methods for producing such polypeptides by recombinant techniques are disclosed. Also disclosed are methods for utilizing receptor polypeptides and polynucleotides in the design of protocols for the treatment of diseases and diagnostic assays for such conditions.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/00959

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C12N15/85 C12N5/10 C07K14/705 C07K16/28
C12Q1/68 A61K38/17 A61K48/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K C12Q A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE GENBANK Accession No. T48852, 8 February 1995 "EST70319 from Stratagene placenta" XP002066898 see the whole document ---	1-12,20
X	DATABASE GENBANK Accession No. G14442, 4 January 1996 "STS SHGC-10055" XP000206689 see the whole document ---	1-3,20
	-/--	

[X] Further documents are listed in the continuation of box C.

[ ] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
4 June 1998	16. 09. 1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Galli, I

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ZANNETTINO A.C.W. ET AL.: "A powerful technique for isolating genes encoding surface antigens using retroviral expression cloning" J. IMMUNOL., vol. 156, no. 2, 15 January 1996, pages 611-620, XP002066893 see abstract	1-21
A	ADAMS M D ET AL: "COMPEMENTARY DNA SEQUENCING: EXPRESSED SEQUENCE TAGS AND HUMAN GENOME PROJECT" SCIENCE, vol. 252, no. 5013, 21 June 1991, pages 1651-1656, XP000645049 cited in the application see abstract	1-21
A	SIMMONS D. & SEED B.: "Isolation of a cDNA encoding CD33, a differentiation antigen of myeloid progenitor cells" J. IMMUNOLOGY, vol. 141, no. 8, 15 October 1988, pages 2797-2800, XP002066894 see abstract see figure 4	1-21
A	CROCKER P.R. ET AL.: "Sialoadhesin, a macrophage sialic acid binding receptor for haemopoietic cells with 17 immunoglobulin-like domains" EMBO J., vol. 13, no. 19, 3 October 1994, pages 4490-4503, XP002066895 see abstract see figure 2	1-21
A	KELM S. ET AL.: "The Sialoadhesins: a family of sialic acid-dependent cellular recognition molecules within the immunoglobulin superfamily" GLYCOCONJ. J., vol. 13, no. 6, December 1996, pages 913-926, XP002066896 see the whole document	1-21
A	DATABASE GENBANK Accession No. U71382, 12 November 1996 PATEL N. ET AL.: "OB binding protein - 1" XP002066901 cited in the application see the whole document	1-21

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	DATABASE GENBANK Accession No. AA344713, 21 April 1997 "EST50650, H. sapiens cDNA similar to CD33 antigen" XP002066900 see the whole document ---	1-12,20
P,A	TAKEI Y. ET AL.: "Molecular cloning of a novel gene similar to myeloid antigen CD33 and its specific expression in placenta" CYTOGENET. CELL GENET., vol. 78, 1997, pages 295-300, XP002066897 see abstract see figure 1 -----	1-21

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 98/00959

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Remark: Although claims 13 and 14 are directed to a method of treatment of the human/animal body, and claim 15 (as far as in vivo methods are envisaged) is directed to a diagnostic method, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see further information sheet

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [X] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

see further information sheet, subject 1.

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International Application No. PCT/ US 98/00959

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: (1-21) - partial

An isolated polynucleotide encoding the HMACR70 cell surface receptor (Seq. IDs 1,18), polynucleotides complementary thereto, fragments thereof, homologs thereof.

An expression system comprising said polynucleotides, a process to produce a cell harbouring said expression system. Said cell or membrane thereof expressing said receptor.

A process for the production of the HMACR70 polypeptide comprising culturing said cell.

Said HMACR70 polypeptide or homologs. An antibody directed against said polypeptide.

Therapeutic and diagnostic applications.

A method to identify agonists or antagonists to said receptors. Said agonists and antagonists.

2. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 2,3,19

3. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 4,20

4. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 5,21

5. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 6,22

6. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 7,23

7. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 8,24

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

8. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 9,10,25

9. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 11,26

10. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 12,27

11. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 13,28

12. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 14,29

13. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 15,30

14. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 16,31

15. Claims: (1-21) - partial

Idem as subject matter 1 but limited to Seq. IDs 17,21